

Scientist tests the public trust

An unreported accident in a virus laboratory at Yale is but one of many factors that challenge society's trust in scientists' promises to do hazardous research.

EARLIER this month a scientist at the Arbovirus Research Unit at Yale University was spattered with deadly Sabiá virus when a test tube broke in a high containment facility. Incredibly, he failed to report this clear-cut accident to university authorities. Instead, he went about business as usual, even taking a trip to Boston where he stayed with (and exposed) friends.

The accident was not discovered by authorities until the researcher became ill with fever and confessed what happened. He was put into isolation and his known contacts are under precautionary surveillance. Although there is little likelihood that they, too, will become ill because the rodent-borne virus does not spread easily from one person to another, the accident and the scientist's behaviour are not a trivial matter.

Sabiá virus was first isolated from an office worker, whose route of exposure to the virus remains a mystery, and who died in 1990 in the town of Sabiá, Brazil. Viral samples were eventually sent to Yale, and by Yale to two other World Health Organization "reference" laboratories in the United States: the Centers for Disease Control and Prevention (CDC) in Atlanta and the U. S. Army's virus laboratory at Fort Detrick, Maryland. The virus can cause lethal haemorrhaging and everyone who studies it knows how dangerous it is.

Why the infected scientist failed to report the accident at Yale is not known. But its effect is both known, predictable and potentially more harmful than the virus itself. Federal, state and local health authorities have threatened to close the Yale Arbovirus Unit as a threat to the community. From a public relations standpoint, this is a rational response. After all, who would want to live in the vicinity of a laboratory that houses dangerous viruses if the scientists cannot be trusted to follow the strict rules of containment that they themselves have laid out? And why should the public feel reassured that things will go right in the future when the cause of the present incident—a broken vial of virus—was so unambiguous.

But if the Yale arbovirus laboratory were to be closed (which is unlikely this time) what then? We may all live (or die) to regret it. Viruses like Sabiá, classified as an arena virus, are often described as 'rare'—largely because their natural habitat is in the wilds of South America or Africa. But by now everyone (certainly public health officials) should know that these agents may not be as rare as we like to think, or that, as is posited with the AIDS virus, they may be transmitted to more urban populations as people clear the world's remaining jungles for human habitation.

It is because so little is known about these viruses that laboratories such as the Yale unit, headed by Robert E. Shope, and other high-technology laboratories staffed by scientists dedicated to viral analysis are essential to public health worldwide. Good scientists are well attuned to the threat of emerging microbes. Their structure needs to be identified, their behaviour studied, and samples kept for reference and comparative analysis in much the way data bases have been established for the structure of genes whose function is known and unknown. It is our only hope of learning enough to one day be able to either destroy harmful viruses or develop more effective treatment for those who are infected.

And arboviruses are not the only threat. It is not necessary to read the special literature of virology to know that human beings are under siege. Reading the newspaper will do.

Rodent-borne hantaviruses, which like the Sabiá virus cause haemorrhagic fever, were first identified in 1976. But just last year they made headlines when people died mysteriously in the Four Corners area of the US Southwest where Utah, New Mexico, Arizona and Colorado come together. Of course, once public health epidemiologists got on to the case, it became evident that the hantavirus was at work. The hantavirus made headlines again when, in February 1994, it killed a student in the state of Rhode Island. Virologists from the CDC said at the time that the virus seemed to be the fourth known variant of hantavirus they had seen in only a year's time.

The viruses are alive and well. Scientists' understanding of them remains modest. Why, for instance, are some viruses virulent in human beings while others are benign? The answer to that question alone could yield a lot of practical information.

The case for the careful, persistent study of the microbial world should be self-evident. But it should also be evident to researchers that public confidence is badly shaken when scientists who promise to take every precaution fail in their duty. That is what happened at Yale. That is why public health officials had to at least threaten to close the arbovirus unit. And the danger of that happening is why every researcher, in any laboratory in any nation, who is working at the edges of science with agents that cause disease has a special duty to follow the rules—precisely, to the letter of the scientific law. Yale has a duty to let the public know how it is going to act on the public's behalf once the infected but miscreant researcher has recovered. □



AIDS . . . again

Knowledge that something is dangerous is not necessarily a sufficient inducement to change behaviour.

WHERE the older generation of gay men changed their sexual behaviour in the face of HIV, the younger generation is experiencing an increased incidence of infection among homosexuals. Recent data compiled by epidemiologists at the University of California at San Francisco (UCSF) indicate that one in three men who are homosexual or bisexual will be HIV-positive by the time they are 30 years old.

Simply put, these data say that the deadly AIDS message, which penetrated the lives of older men who have witnessed the deaths of dozens of their friends, has not been passed on to teenagers and young adults. But the virus is being passed on with negligent abandon among young men who knowingly chose to practise unsafe sex. According to reports of work by Dennis H. Osmond of UCSF and colleagues, even though these young people (many of whom have been to AIDS counselling sessions) know they are at risk, knowledge alone does not change their behaviour.

These troublesome data merely confirm the weak -to-non-existent link between what one might call 'book-learning' or knowledge from lectures and behaviour. Recent studies of homosexual and heterosexual couples in Europe, in which one partner was known by the other to be HIV-positive, showed that 52 per cent had unprotected sex nonetheless (see *Nature* 370, 400; 1994).

What is it that makes people take such serious risks? Is it youth — simple as that? Or are there other factors particular to the issue of safe or unsafe sex and AIDS? The answers to date amount to little more than pop-psychology, which is not enough. □

Clinton's indecision

US president Bill Clinton's "policy" on the Cuban refugee crisis speaks volumes about his presidency.

As the world watches in horror, thousands of Cubans, fleeing their homeland in search of a better life, are crowding into rafts and putting out to sea in search of America. But America, under policy set forth by US president Bill Clinton, is turning them away. The president's reasoning, which is a bit difficult for a sensible person to follow, is that he will not allow Fidel Castro to dictate US immigration policy.

The city of Miami, Florida, is already a haven overflowing with Cuban refugees who have transformed the cultural identity of the place. And some of the very people who were allowed to immigrate in years past are now taking a 'not-in-my-backyard' stance with regard to their former fellows. 'There is not enough room for more of us', some are saying in effect, giving the president ammunition for his no-admissions policy.

As a half-way measure, Mr Clinton has adopted one of the

least humane and least effective policies he could have. He has stationed US ships in the ocean midway between Cuba and Florida, with instructions to intercept refugees and return them where? To Cuba. Not to their homeland, but to a US base at Guantanamo Bay where they are massed together in crowded tent cities.

So — what does this say about Mr Clinton? That he seems to lack the capacity to anticipate the future. As long as two years ago, scholars were predicting precisely the quagmire that the Clinton administration is now in. A year ago, the Central Intelligence Agency was issuing classified warnings to the White House. And officials at the State Department were also fully cognizant of a looming international relations disaster with Cuba. Nor has Mr Clinton shown himself particularly adept at dealing with the predictable now that it has occurred.

His ability to deal with the crisis in Haiti is equally dismal. There, too, expensive, disease-ridden tent cities erected by the US military sometimes make it difficult to tell the difference between the Caribbean and the sorrowful state of refugees in Africa.

Mr Clinton has taken a middle road to disaster by neither fully admitting nor definitively rejecting the refugees probably heading for US shores. It is a telling statement about his leadership in international affairs. □

Is health reform dead?

Mr Clinton appears ready to compromise on health care reform despite previous denials.

THE comprehensive and deceptively costly program for health-care reform that US president Bill Clinton, Hillary Rodham Clinton, and consultant Ira Magaziner expected to sweep the public off their feet appears dead in the water, at least for this year. Neither Congress nor the US public has been able to fully grasp all of the ins and outs of the 1,342-page plan that the Clinton team has been utterly unable to explain. That may be just as well because the dead plan was so full of intrusive provisions that no man, woman, child, or business in the country would have escaped the heavy hand of a well-meaning administration whose ideas for reform violate the Hippocratic oath to do no harm.

Individuals feared (rightly) that their choice of doctor would be limited. Small businesses said they cannot afford to pay workers' insurance premiums. Some politicians said the plan was too liberal; others said it was not inclusive enough. And no one could agree how to pay for it. In short, the Clinton plan (with its innumerable complexities) was an easy target, shot down often for the right reasons. It promised comprehensive benefits to everyone but, once its true dollar costs were revealed, looked too expensive even for a wealthy nation.

One irony of the latest news about health reform's demise is this: on the very day most observers declared that the Clintons' way would break the bank, surgeons in California performed open heart surgery on a young orangutan — a member of an endangered species. □

Japanese science agencies stress networks in request for budget hike

Tokyo. Japan's Science and Technology Agency (STA) and the Ministry of International Trade and Industry (MITI) hope to create a more open government research system through 1995 budget requests submitted to the Ministry of Finance this week. The agencies say their budget plans will encourage closer links between national research institutes, universities and industry.

Both MITI and STA plan to put more money into computer networks that will link government research institutes and universities and to invest in new facilities and programmes to encourage cooperative research between different branches of the government and industry.

Despite a severe squeeze on government finances due to plummeting tax revenues during the recession, both MITI and STA have made ambitious budget requests that are between 8 and 9 per cent more than this year (see table). The requests will be trimmed back in negotiations with the Ministry of Finance before finalization at the end of December.

STA hopes almost to double its spending

Rocket success brings sighs of relief

Tokyo. Officials of Japan's Science and Technology Agency, who are requesting a large increase in the budget for the agency's space programme (see above), were able to breathe a sigh of relief Sunday (28 August) when Japan's new H-II rocket finally lifted off from the Tanegashima space centre, after two aborted launch attempts (see *Nature* 370, 586; 1994).

The H-II, which has a sophisticated liquid fuel engine in the first stage similar to that of the US space shuttle, carried a 2-tonne test communications satellite. If all goes well, the satellite will move into geostationary orbit after firing its engines three times over the coming weeks.

The new rocket is vital to the agency's future space programme. Most of the proposed increase in the agency's space budget for 1995 will be spent on large Earth observation satellites to be launched by the H-II, and on development of an unmanned space shuttle called HOPE that will be carried on an upgraded version of the rocket. **D. S.**

Highlights of Japan's R & D Budget Requests

(in billion yen US\$1 = ¥100)

MITI	1995	% change
Total R & D budget request	307.1	+7.6
Japan Key Technology Center	27.0	+3.8
Industrial Scientific Technology	25.6	+8.5
New Sunshine Project	55.8	+5.7
Real World Computing Project	6.0	+20.9
Intelligent Manufacturing System Project	1.3	+1.8
Human Frontier Science Programme	3.6*	0.0
Global Environment Research	13.1	+9.1
Space Flyer Unit	8.8	+5.2
Research Information Center (new)	0.6	—
NEDO fellows (new)	0.5	—

* includes 2.1 billion yen from Science and Technology Agency

Science and Technology Agency	1995	% change
Total R & D budget request	661.0	+9.2
Special Promotion Funds	19.0	+22.6
Space	183.7	+8.7
Nuclear Power	346.0	+6.8
ITER	8.3	+3.7
Interministry computer network	2.0	+81.8
Human Genome	2.5	+31.6
SPRING-8	15.7	+42.7
Frontier Research System	3.9	+19.9
STA Fellowships	2.2	+15.8

on an interministry computer network now under construction that will link government research institutes and universities via a high-capacity backbone network with access nodes in Tsukuba science city, Tokyo and the Kansai region encompassing Osaka and Kyoto. The agency also requests funds for establishment of several new databases related to the network.

Similarly, MITI has a request of ¥600 million (US\$6 million) for a new centre that will provide users in universities and private industry with access to the ministry's powerful supercomputer centre in Tsukuba science city when they carry out joint research with the ministry's research institutes.

MITI's Agency of Industrial Science and Technology already has a Research Information Processing System (RIPS) that links all 15 of the agency's research institutes around Japan to the supercomputer centre in Tsukuba. But RIPS can be used only by the agency's 2,500 researchers. The aim of the new centre, which will be based in Tsukuba, is to open up the computer system and databases to outside users.

Linked to this initiative is a request to establish a new fellowship scheme to be run by the New Energy and Industrial Technology Development Organization (NEDO), a semi-government organization under the ministry that administers many of MITI's research and development projects. NEDO will support about 60 fellows a year drawn from universities, private companies, na-

tional research institutes and from overseas to participate in NEDO projects, including new projects starting next year for the development of computer networks and research information systems.

Traditionally, there has been very little communication between government research organizations belonging to different ministries and agencies because of the bureaucratic walls that separate them. For years the government has been talking about the need for more interaction, but only recently have concrete schemes emerged. If the plans by MITI and STA to establish a more open system are fully realized, this could in the long run revolutionize the way research is carried out and administered in Japan.

MITI has also asked for a 21 per cent increase for its Real World Computing (RWC) project, a 10-year initiative by MITI institutes and private industry to develop new massively parallel computers. The ministry has requested ¥6 billion for the project next year and plans to invest a total of about ¥70 billion by 2002.

The project, which began in 1992 on a small budget of ¥0.9 billion, is a follow-up to Japan's famous (but not very successful) fifth-generation computer project. Unlike that project, however, which was carried out in a central laboratory, the RWC project is distributed among numerous laboratories in Japan and one or two in Europe that are (or will be) linked by a high-capacity computer network to allow remote operation of massively parallel systems under development.

The STA, meanwhile, plans to expand the human genome project to ¥2.5 billion, 30 per cent more than this year. The agency's researchers have until now concentrated on mapping of the human genome and the establishment of a gene bank at the Tsukuba Life Science Center of the Institute of Physical and Chemical Research, as well as on the development of sequencing technology. But from next year, the project will move towards full-scale sequencing of the genome, says an agency official.

The human genome project has been one of the driving forces behind the establishment of computer networks in Japan. And human genome researchers should be among the first to benefit from the new interministry network and the hoped-for improvement in links between different branches of the government. **David Swinbanks**

'Drop-out' row hits soft spot for UK universities

London. A row over disputed university 'drop-out' rates published last week has highlighted the problems faced by science, engineering and technology students, who are more likely to fail than their peers studying arts and social sciences.

Figures published in *The PUSH Guide to Which University*, a rough guide aimed at helping prospective students to choose a university, show that in some UK universities as many as 1 in 5 students fail to graduate. According to the figures, the average drop-out rate for students expected to graduate in 1992 was 13.2 per cent.

The PUSH 'league tables', which are calculated using statistics from the Committee of Vice-Chancellors and Principals (CVCP) and the Universities Funding Council (UFC), have been attacked as inaccurate by some of the universities singled out. But even if some of the figures are wrong, the guide has succeeded in drawing attention to a mounting problem which many in the universities would rather keep under wraps.

An earlier study of Scottish universities confirmed anecdotal evidence that the highest non-completion rates are in engineering and other subjects with higher-than-average numbers of hours spent in the laboratory or lecture room. The study was funded by the Scottish Office Education Department to find out why the rate was 4 per cent higher in Scotland than in the rest of the United Kingdom.

Robin Knops, professor of mathematics and vice principal of Heriot-Watt University in Edinburgh says that his university has taken steps to address the problem since the Scottish Office report came out. He thinks that the difficult ideas and concepts in-

volved in science, engineering and technology subjects have combined with an overloaded syllabus and all the normal pressures of student life to aggravate the problem.

The PUSH guide lists Kings College London, part of the University of London, as having a drop-out rate of 21.3 per cent, second highest in the 'league'. But the college says the correct figure is 14 per cent.

A spokesman for the London School of

senior assistant registrar, says the university's unusual mix of courses tends to inflate its non-completion rate. Salford runs a high proportion of '2+2' courses, where students with one A-Level can study for two years at an associated college for a diploma before deciding whether to study for a degree. Although all are registered at Salford as degree students, many successfully complete the first two years but decide not to continue. Once these courses are taken into account, Salford calculates its non-completion rate as 16 per cent compared to the 26.1 per cent quoted by PUSH.

Others, such as Brunel University on the outskirts of London and Heriot-Watt, accept that the PUSH figures of 21 per cent and 20.3 per cent respectively are about right. A Brunel spokesman says the drop-out rate is a "matter of concern" and that it is conducting an internal analysis of the reasons. Both Brunel and Heriot-Watt point to the very high proportion of students studying science, technology and engineering subjects at their institutions.

According to the latest edition of the CVCP/UFC management statistics, the percentage of successful leavers is 86 per cent for physical sciences and 83 per cent for engineering and technology. That compares with 90 per cent for social studies, humanities and business administration.

Knops is confident that the measures introduced at Heriot-Watt can stem the tide of non-completions. They include a compulsory mentor scheme for each student, remedial teaching and support for any student identified as struggling and a modular course structure with students being examined at the end of each module rather than at the end of the year.

Maggie Verrall

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

More students in — fewer out?

Economics (LSE) says that the college is "very, very annoyed" that PUSH has released figures into the public domain that the LSE says are wrong. PUSH calculates drop-out rates for LSE to be 19.4 per cent, but LSE says the true rate was 7.9 per cent in 1992.

According to the CVCP, which commissions the Universities Statistical Record to compile statistics annually, using them to create league tables is likely to be "misleading" because of the different subject and student mixes of the institutions. A CVCP spokesman said that the national success rate of 87 per cent for 1992-93 was something that the country "should be proud of".

At the University of Salford, Les Kilby,

State of Washington plans cull to save endangered steel-

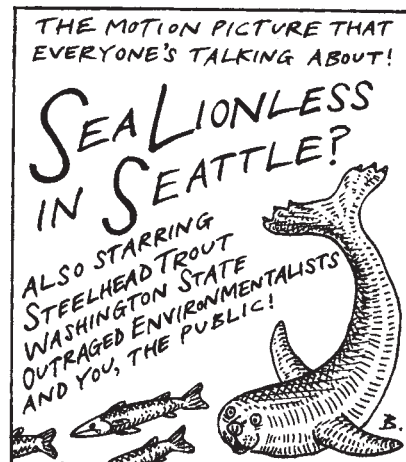
San Francisco. Environmentalists and government officials in the northwest United States are at odds over a plan to protect an endangered breed of fish by killing off the highly popular sea lions that feed on them.

The state of Washington has asked the federal government for permission to "fatally remove" about half-a-dozen California sea lions that are ambushing wild steelhead trout on their way to spawn. Both species are protected by federal law.

In a letter to the Secretary of Commerce, Ron Brown, the state has called for a task force to rule whether killing the sea lions is justified — a first step the state must take, under law, before going ahead with the cull. Environmentalists hope that a less drastic solution can be found. "We are trying to find ways to make the fish less vulnerable," says Julia Reitan of the Seattle section of the Sierra

Club, an environmental pressure group.

That something needs to be done is indisputable: last year only about 70 of the



steelhead, a popular and beautiful food fish, reached their spawning ground. This contrasts with 2,965 known fish in 1986 and is far below what is needed to sustain the species.

The fish, an anadromous form of rainbow trout, swim from the salt water of Puget Sound to Lake Washington on their way to spawn in the Cedar River. Their way is blocked by the Ballard Locks, where the waterway narrows. Sometime around 1980, a pod of approximately 60 sea lions began waylaying the steelheads at the locks. Of those sea lions, about five or six are particularly accomplished hunters — and these are the target of the proposed cull.

Unlike most acts of depredation, which usually take place in remote areas, the sea lions attack the steelhead in downtown Seattle, where the spectacle has become a popular tourist attraction.

Radioastronomers hope for world observatory

Washington. The cost of future radio telescopes implies that an international consortium will be needed to build them, astronomers say. But several large obstacles will have to be overcome, particularly if the richest countries — the United States and Japan — are to be involved in such a collaboration.

Ideally, a World Radio Astronomy Observatory (WRAO) would operate along the lines of the US National Radio Astronomy Observatory, with a small central bureaucracy overseeing a facility open to all astronomers from anywhere in the world.

But the political problems are daunting. Most national governments would expect their scientists to get a guaranteed fraction of time at a telescope, in return for their financial contribution. This is not the way to get the best science from the facility — but few countries are rich enough to share the United States view, that access should be based on scientific merit alone. This causes friction whenever the US is involved, because American scientists take for granted that every facility will be open. And if Japan is involved, extra costs associated with the Japanese method of doing big science become a problem.

At last month's conference on "Radio astronomy visions for the 21st century" at Penticton, British Columbia, astronomers gathered to ponder which big problems remain in their discipline, and what instruments they need most.

They agreed that the most interesting problems are "when and how do galaxies form?", and "what does their early evolution look like?" But the answers to these questions will not come cheap. The next generation of radio telescopes will cost \$150–\$300 million to build, and \$5–\$15 million

per year to operate. This led Ron Ekers, director of the Australia Telescope National Facility, to suggest that it was time to consider founding — and funding — a World Radio Astronomy Observatory.

Two projects widely acclaimed at Penticton could potentially fit the bill. Roy Booth, director of the Onsala Space Ob-



The \$150 million millimetre-wave array: under consideration at NSF.

servatory in Sweden, proposes an array of forty 15-metre — or ninety 10-metre — telescopes operating as an interferometer at wavelengths from 0.8 mm to 3 mm, and located in Chile. Booth says that with current technology this would cost \$250 million to construct, but hopes that better antenna design would bring the price down to \$150 million.

Robert Braun, of the Netherlands Foundation for Research in Astronomy, suggests a phased-array telescope with a total collecting area of 1 km², operating at wavelengths from about 21 cm to 1.5 metre, and primarily studying emission from atomic hydrogen. The detectors, if located in the Netherlands, would be distributed over an ellipse 30 × 50 km in size, with a few outlying telescopes 150 km from the centre

and a densely packed region in the inner 400 metre. This would combine very high resolving power with high sensitivity, he says. The estimated cost today would be \$300 million, but Braun hopes this could come down to \$100 million with advances in technology.

Both of these ambitious projects would be able to study emission from gas in galaxies very early in the history of the Universe. Booth's telescope would study molecular gas from which stars form — a new but hot branch of astronomy. Braun's telescope would study atomic hydrogen in the clouds that condense to become galaxies.

Both proposals meet the rough guideline that any new telescope should realize a factor-of-ten improvement in sensitivity to be worth building. Booth's would have ten times the collecting area of the 30-metre telescope at Pico Veleta, Spain, and Braun's telescope would have 75 times the collecting area of the Very Large Array in New Mexico, and 14 times the collecting area of the largest single-dish radiotelescope in the world at Arecibo, Puerto Rico.

It is unlikely that Sweden or the Netherlands could fund these projects alone. Realistically, their funding will involve either many countries joining together to fund a World Radio Astronomy Observatory, or getting the Americans or Japanese to foot most of the bill.

In the United States, the National Science Foundation is already considering a proposal from the National Radio Astronomy Observatory to build the \$150 million millimetre-wave array (MMA) — a smaller and less powerful version of Booth's proposal. The United States is unlikely both to build the long-awaited MMA and to be a major partner in an international collaboration.

Japanese astronomers also have their own plans for a millimetre-wave array, but, in the tradition of Japanese science, it will cost more than other countries may be willing to pay. M. Ishiguro, of the Nobeyama Radio Observatory, estimates a cost of about \$1 billion for an array similar to the MMA, but operating to wavelengths as short as 0.375 mm. Japanese radioastronomers are confident that they will have \$300–\$400 million to spend on their project near the end of the century, but will be hard pushed to attract others to contribute \$600–\$700 million, much of which would go to Japanese industry for development costs.

Last month, the general assembly of the International Astronomical Union in The Hague voted to set up a committee to investigate the best options for building a millimetre-wavelength array. The committee will be hard put to reconcile the conflicting interests of the governments that would pay for such an array and the scientists who would use it.

Leslie Sage

head trout from hungry sea lions

Efforts to control the fish kill have gone on for years. State fish and game authorities have tried everything from throwing firecrackers at the sea lions and playing tape-recorded *orca* sounds to removing some sea lions as far away as the Channel Islands off Santa Barbara, California. Most of the animals so displaced found their way back to Seattle, more than 1,000 miles up the coast.

Environmentalists claim that the sea lions are only one cause of the decimation of the steelhead, pointing out the degradation of the streams beyond the lakes, pollution and the rapid urbanization of the area. Reitan says that local Indians who fish for the trout claim the fish ladders at the locks were badly designed and that rebuilding them could make it possible for the steelheads to survive. And Bruce Sanford of the state Department of Fish and Wildlife says it is possible that if they removed the most voracious

hunters, other sea lions might just improve their skills and replace them.

So why does the state want to kill the most aggressive animals? Local officials say that it takes several hundred fish to sustain the gene pool and even if the sea lions are only one cause of the decline, their appetites alone may wipe out the entire steelhead population. "It's gone beyond critical," Sanford says.

Unlike its close relative the salmon, which is also declining dramatically in the area, the steelhead spawns more than once and each year another brood tries the locks. Just because only 70 fish made it last time does not mean the fish is doomed, Sanford says. The task force's job is to determine whether evening the odds at the Ballard Locks will give the fish a chance to renew their population. A decision will be made at the end of the year.

Joel Shurkin

Hungarian coalition has pro-science leanings

Munich. Science has not escaped entirely from Hungary's new belt-tightening programme — but it fared better than most other sectors of the economy in the revised 1994 budget announced by the new government last week.

The coalition of socialists (reformed communists) and free democrats which formed a government in July has inherited a severe budget deficit, and the cabinet has spent the summer discussing how it can claw back some HUF17 billion (around US\$150 million) from this year's budget.

Its conclusions — to be ratified by parliament later this month — together with the new government's declared interest in building a more robust structure for science policy, indicate that this government may be more sympathetic towards science than its predecessor.

Negotiations with the Hungarian Academy of Sciences persuaded the socialist prime minister, Jyula Horn, that the academy, which runs more than 40 research institutes in Hungary and also acts as a learned society, should be considered a special case because of the severity of its

treatment over recent years.

"It is not possible for the government to look only at one year and make the right decision," argues László Keviczky, general secretary of the academy, whose budget has more than halved in real terms since 1992. The academy has also had to deal with perpetual political hostility — including charges of harbouring communist sympathies — despite the extensive reforms it has undergone. After years of fighting for legal status, the academy finally achieved it last spring. "We have cut 35 per cent of our staff [since 1989] — the highest percentage cut of any state-owned sector," says Keviczky.

The government has been responsive to these arguments, and is now asking the academy to sacrifice only HUF120 million from its 1994

budget of around HUF5 billion. Keviczky sees this as a strong sign that the new government is sympathetic towards science. But he is under no illusion that it will be easy

to hold its ground when the 1995 budget is set in late autumn. The academy comes under the aegis of the ministry of culture and education, whose new minister, the 31-year-old free democrat Gábor Fodor is an unknown quantity as far as his position on science goes.

Applied science did not fare so well last week, suffering severe cuts comparable with other, non-science sectors. The National Committee for Technological Development (known as the OMFB), which is responsible for allocating grants for applied research, will lose HUF600 million from its

1994 HUF4.6 billion budget. Lajos Myiri, deputy director of OMFB, says that savings will probably be made in its infrastructure funds rather than in its competitive research awards, which it is determined to defend at all costs.

The government is also considering new ways of organizing Hungary's research policy now that the dust is starting to settle on post-communist reforms.

The OMFB, whose president, Ernő Pungor, was a minister without portfolio in the previous government, will now report to the minister of industry, socialist László Pál. This move is intended to streamline government, but it almost certainly also has a political intent. Although it will no longer have direct contact with the cabinet, OMFB will nonetheless maintain its current structure and functions. The agency is also responsible for international relations in both basic and applied science, and advises the government on research policy.

The government is considering setting up a science and technology committee at cabinet level, which would comprise the ministers for industry and education, and representatives from the academy, OMFB, the independent grant-giving agency OTKA and other scientific bodies.

The prime minister has welcomed the idea in principle, and, says Keviczky, has indicated that he might be prepared to act as chairman, should the committee be established. Such a committee might be similar to the Czech republic's new government advisory committee on science and technology, whose round-table meetings include scientific advisers as well as the finance and economics ministers (see *Nature* 368, 386; 1994).

Allison Abbott



Young, gifted and left-of-centre: Gábor Fodor, Hungary's fourth most popular politician, is now in charge of basic science.

Tom Stoddart (Nat. Pictures)

Euroagency for drug evaluation sets up shop in London's East End

London. The new European Medicines Evaluation Agency (EMA) sets up home this week in the Canary Wharf development in London's East End, with an initial staff of around 200.

It will be headed by Fernand Sauer, former head of pharmaceuticals at the European Commission. The agency will coordinate licensing of medicines in all 12 member states of the European Union (EU). At the moment companies must apply to each state individually: under the new system, a medicine whose safety, efficacy and production standards are approved by the EMA may be sold in all EU countries.

Some biotechnology products will be obliged to go through the EMA instead of national agencies. For other products the new route will be optional, but the advantages are clear — products should reach a bigger market more quickly. The EMA hopes to cut licensing time from several years to a maximum of 12 months.

But drug companies are reserving judgement on the new system. The EMA will delegate assessment of particular medicines to a national agency that will act as rapporteur. A second national agency will act as co-rapporteur, to provide a "double-check". Bob Blakie, head of British Biotechnology, says that the role of these rapporteurs remains unclear. And as companies will not have a final choice of rapporteur country (they are invited to make suggestions) they may lose the advantage of having dialogue with the relevant regulatory agency during product development.

A. A.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Court case blows away US–French truce on industrial espionage allegations

Paris. Allegations by the US company Texas Instruments (TI) that technology stolen by French spies ended up on a patent awarded to the computer company Bull have reopened the simmering row between the United States and France on industrial espionage.

The French state-owned computer company fired the first shot in this latest dispute — revealed last week by the French magazine *L'Express* — by suing TI for infringing a 1983 Bull patent for a chip designed to process confidential information. But TI's return fire sent the patent lawyers diving for cover: the patent, it said, was based on TI technology "unlawfully expropriated by one or more agents of the French government covertly placed within TI... to obtain technical information for Compagnie des Machines Bull".

Bull seems to be asking for a thrashing. Even if its claims about the 1983 patent are valid, TI is well positioned to support its allegations of spying. In 1989 the US Federal Bureau of Investigation (FBI) and the Central Intelligence Agency (CIA) identified a senior TI employee as a French spy, confirms Maurice Botbol, publisher of the Paris-based "Intelligence Newsletter".

The employee, identified only by the code-name "MD", was part of an elaborate programme set up by France following the election of François Mitterrand as president in 1981. Pierre Marion, then head of the French external security service, sought out industrial and technological targets, and

began recruiting large numbers of "honourable correspondents" within foreign companies. The practice then — and probably now, he has said — was to pass on information to French companies.

The French and US governments will be less concerned by who owns the patent than by the risk that the dispute could undermine their agreement not to accuse each other of spying in public, says Botbol. Under the terms of this partial truce — intended to prevent souring of diplomatic relations, and keep the way clear for cooperation against terrorism and nuclear proliferation — neither side has agreed to stop economic spying, he says, but they have agreed not to go beyond what is "tolerable".

The arrest of the TI employee along with a network of French spies implanted in other US technology companies in 1989 was quickly followed by a meeting between senior US intelligence officials and the head of the French secret service, Claude Silberzahn. France, caught red-handed, owned up, promised not to do it again, and managed to avoid any real sanction.

Last year, secret services from the two countries signed a fresh "non-aggression" pact. This followed French dismay after the US company Hughes Aircraft, acting on information from the CIA, withdrew from the Paris Air Show because of the threat of espionage. French observers allege that the CIA fabricated the threat partly to discredit the air show, and partly to persuade Congress that the CIA was still worth

funding, despite the collapse of the Soviet threat.

Truce or no truce, TI seems determined to fight off Bull's patent challenge in the courts. According to *L'Express*, the company has already asked Bull to supply documents which it claims reveal its relations with the French secret service, and would demand that "MD" stand as a witness. Neither company is willing to comment either on the court case or the background to it. But observers predict that somehow the matter will be taken down a long dark alley and quietly smothered.

Industrial espionage now accounts for two-thirds of all spying, according to the French counter-espionage service. IBM Europe describes espionage as a "continuing problem". One US agency has estimated that such spying costs the US economy \$100 billion annually, while the figure in France has been estimated at FF25 billion.

More prosaically, some French companies complain that the intelligence services are not collecting enough of the right sort of information, and that it is not equitably distributed between them. For their part, the counter-espionage agencies are campaigning to persuade French companies and laboratories to take more precautions. Even the French biomedical research organization INSERM recently warned its researchers of the growing risks of espionage arising from the internationalization of science and its growing strategic importance (see *Nature* 367, 587; 1994).

Declan Butler

Wyden widens net of probe into drug marketing

San Francisco. A subcommittee of the United States Congress is widening its investigation into allegedly fraudulent marketing activities at the Californian biotechnology company Genentech to include Eli Lilly and other pharmaceutical companies.

Three federal agencies and the House of Representatives subcommittee on small business regulation chaired by Ron Wyden (Democrat, Oregon) have been investigating the marketing and promotion of Protopin, Genentech's human growth hormone.

Steve Jennings, staff director for the subcommittee, says the lawmakers are asking the agencies to look at "unsettling evidence" that marketing and research improprieties are common throughout the industry. Jennings thinks that new drug marketing techniques blur the lines between marketing and research: research grants, direct reimbursements for prescribing doctors and cash in-

centives are all problems, he says. The committee has written to Genentech and Eli Lilly requesting information about their research into human growth hormone.

On 23 August, the Inspector General of the Food and Drug Administration issued a special fraud alert warning that many drug companies are illegally going beyond traditional advertising and education. They are illegally offering valuable benefits to physicians, suppliers and, increasingly, patients, in exchange for selecting specific prescription drug brands, the notice said. The warning mentions not only prizes and cash or gifts offered for prescriptions, but also grants to physicians and clinicians for studies of questionable scientific merit.

In the case of Genentech, regulators are inquiring into alleged bribes for prescribing its human growth hormone; marketing for short stature; and promotional activities by a nonprofit foundation that

Genentech helps to fund.

This week the Foundation on Economic Trends, a pressure group led by Jeremy Rifkin, called on the Drug Enforcement Agency to prosecute doctors who prescribe human growth hormone to children who are not growth-hormone deficient. The group based its request on a law that prohibits sale of the hormone for anything other than the treatment of a recognized medical condition.

Rifkin also contends that Genentech both directly and indirectly funded height-screening programmes in hospitals and four US school districts to help draw new patients. Genentech defended its efforts to fund research on short-stature children. "All doctors are highly trained, highly intelligent people who are able to make judgements on their own," says a company spokesman, adding that pharmaceutical funding would not wash their integrity away.

Sally Lehman

Lack of political will, clinical precision stalls RSI research

London. The British government is under growing pressure to fund more research into repetitive strain injury (RSI), the ill-defined condition afflicting keyboard users and other workers which trade unions say already costs UK industry £400 million a year in lost working time.

Occupational health experts say that the lack of money for research into the mystery ailment stems partly from the lack of a precise clinical definition, and partly from buck-passing between government departments about who should take responsibility.

Critics also charge that the government is siding with business interests which fear compensation costs arising from greater recognition of RSI. Many employers have been slow to recognize the syndrome, and drew encouragement when a judge presiding over a compensation case last year said that the label RSI was "in reality meaningless".

But a survey of 53 leading physicians specializing in occupational medicine has found that a better scientific understanding of RSI would help them to attack the problem. "Most people concerned with occupational health look upon musculoskeletal disorders as a major research priority," says Malcolm Harrington, professor of occupational medicine at the University of Birmingham's Institute of Occupational Health, who designed the survey.

Very little money is finding its way into RSI research. In June William Waldegrave — then UK science minister — told parliament that the Medical Research Council (MRC) was not funding any at all. In answer to another parliamentary question, the Department of Health said RSI was the responsibility of the Department of Employment.

Ron Mulelly, honorary secretary of the RSI Association — the main support group for sufferers — says the health department's attitude is "lamentable". "We are concerned with the medical problem," he says, "and if that isn't a matter for the Department of Health, I don't know what is."

A spokeswoman for the MRC says that all applications for grants are assessed on their individual merit, and that the council has no specified areas that it will or will not fund. But Peter Buckle, head of the ergonomics research unit at the Robens Institute of Health and Safety in Guildford says that, after making soundings at the MRC about a research proposal, he came to the conclusion that an application for RSI-related research would not be welcomed.

The growth of RSI was highlighted last month when the Compaq company said it would put warning labels on computer keyboards directing users to advice on how to avoid developing hand and wrist injuries.

Occupational health specialists say they need research to establish a more precise clinical definition of the illness. The current confusion, they say, makes it hard for sufferers to get recognition of their condition, proper treatment, sickness benefit or compensation in the courts.

Many bodies — including the Health and Safety Executive (HSE), the government

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Carpal-tunnel syndrome: better defined.

workplace health and safety agency — prefer to use the umbrella term 'work-related upper limb disorders'. This encompasses accepted medical conditions such as tenosynovitis and carpal-tunnel syndrome, as well as the less well-defined but more familiar RSI.

Owen Tudor, national organizer of the Trades Union Congress's "Don't Suffer in Silence!" campaign, is calling for urgent action from the health department. "With 200,000 sufferers a year, the Department of Health should set up a whole research programme," says Tudor.

The Robens Institute has had some funding for RSI research from the HSE, and has been looking to industry for further support. But it finds that private sector employers fear that to fund RSI work on their own might be seen as an admission that the work they are asking employees to do could cause RSI.

Insurance companies, which are picking up much of the illness's associated costs and have seen related claims double in the past five years, have funded some research in the past. Suzanne Moore, a spokeswoman for the Association of British Insurers, says they may pay for more in the future.

But Harrington says that the best way forward would be to secure funding by industry through the British Occupational Health Research Foundation. This body was set up three years ago partly in response to cuts in government spending on occupational health. Harrington also stresses that, although prevention through ergonomic evaluation is crucial, it may not be possible to "engineer the whole thing away", and that the main need is to establish the pathogenesis of RSI and related complaints.

Maggie Verrall

India moves gently to liberalize drug industry

New Delhi. The Indian government is liberalizing the pharmaceutical industry, lifting price controls on several drugs and allowing greater foreign participation in drug companies.

Officials say the policy to be announced this week will help to make quality drugs available at fair prices, boost research and create a good climate for investment. But critics charge that it will push drug prices beyond the reach of the poor, and lead to the takeover of Indian drug companies by multinationals.

Current price controls cover drugs whose annual sales exceed Rs5 million (\$160,000). Under the new policy, this has been increased to Rs40 million. As a result, the price of 73 drugs will be controlled compared to 143 before. Foreign companies previously limited to stakes of 40 per cent in Indian affiliates can now take 51 per cent of companies making bulk drugs — and more than 51 per cent of companies making drugs with recombinant DNA technology.

The policy stops short of industry's demand for complete decontrol, says Anant Thakore, president of the Indian Drugs Manufacturers Association. "It is two steps forward, one step backward," he says, arguing that there is no need for any price controls. But according to Mira Shiva of the Voluntary Health Association of India, prices of some "uncontrolled" drugs registered a 100 per cent rise during 1990–92.

Prices of essential drugs, as listed by the World Health Organization, and those required for national health programmes will continue to be controlled. Price control will also cover drugs that are a monopoly of one manufacturer. A national pharmaceutical pricing authority will be set up to update the list of controlled drugs and fix their price, and a new national drugs authority in the health ministry will monitor drug quality.

The government hopes the new policy will encourage Indian companies to invest in research through tax benefits and duty concessions. Indigenously developed drugs and diagnostics will not be subject to price control for at least 10 years, for example. But Thakore says the 6–7 per cent profit the industry makes "is hardly enough to make the investments in research needed by us to survive in the post-GATT regime".

The new policy, together with expected changes in Indian patent law, may also encourage foreign companies to invest in Indian research, taking advantage of low salaries for scientists. At present India recognizes only process patents and not product patents, but under GATT (the General Agreement on Tariffs and Trade) it has pledged to accept both. **K. S. Jayaraman**

The National Medical Store Bank

Botanical collections at risk

SIR — We wish to draw your readers' attention to the outcome of an important conference held by UNESCO and the International Council of Scientific Unions (ICSU) at the Komarov Botanical Institute in St Petersburg in December 1993. Among other things, the status and the future of botanical collections in the former Soviet Union were discussed.

There are more than 25 million plant specimens, living and preserved, in more than 230 collections in the former Soviet Union. These constitute a botanical resource of incalculable international importance. Because of chronic underfunding, these collections are at immediate risk of irreparable damage from unsuitable and deteriorating storage conditions such as inadequate pest control and fire precautions. Use of the collections is hampered by insufficient staff (some have no staff at all), inadequate curation and the impediments to the loan and exchange of specimens caused by an expensive and unreliable postal system. There is little hope of an improvement of state funding in the near future. Staff morale is low.

These problems must be of great concern to the international botanical community. Apart from money, the collections are urgently in need of the skills necessary for planning and ensuring their long-term survival. The conference recommended the following actions:

- The appropriate authorities should assume full responsibility for collections on their territories.

- UNESCO should plan an international programme in the field, and should declare an official Year of Natural History Collections to that end.

- The public, governments and international institutions should be made aware of the importance of the collections; they and donor agencies should develop programmes to support the collections.

- The Komarov Botanical Institute should be designated a part of the Russian National Heritage.

- Botanical institutes within the former Soviet Union should produce inventories of their collections and make the information available internationally.

- Each collection should develop a statement of its goals and a realistic plan for achieving them.

- International collaboration should be strengthened by the improved exchange of material and information as well as by staff training in herbarium management, through the *Panarctica Biota* project and

the development of computerized databases.

In our opinion, the institutes responsible for botanical collections in the Soviet Union need not only financial assistance but also help in formulating and executing plans for the future. They need to learn how to fight for their own survival.

D. S. Ingram

Royal Botanic Garden,
20A Inverleith Row,
Edinburgh EH3 5LR, UK

G. Li. Lucas

The Herbarium,
Royal Botanic Gardens, Kew,
Richmond, Surrey TW9 3AE, UK

R. Huxley

Department of Botany,
Natural History Museum,
Cromwell Road, London SW7 5BD, UK

Medicinal plants

SIR — B. K. Holland's Commentary, "Prospecting drugs from ancient texts"¹ proposes a re-examination of Graeco-Roman pharmacological literature to screen for drugs of potential therapeutic use. This important proposal raises three problems that deserve consideration.

First, a philological problem. Although some aspects of Graeco-Roman pharmacology have been studied in depth — for example, its relationship to folklore on the one hand and literate learning on the other² — technical texts such as collections of recipes and drug handbooks have attracted much less attention from philologists and historians of medicine. As a result, only a limited number of old (and sometimes outdated) critical editions of these texts are available. This is true not only of writers considered, rightly or wrongly, to be of secondary importance (for example, Marcellus of Bordeaux³ and Theodorus Priscianus⁴), but also of acknowledged masters of ancient botany and medicine such as Dioscorides and Galen. For example, the reader interested in Dioscorides' *Euporista*, a short handbook that classifies hundreds of medicinal plants on the basis of their therapeutic usage, must still depend on the edition published by C. G. Kuhn more than a century and a half ago (1830).

It is hardly surprising, then, that many anonymous *libri receptarii*, lists of drug prescriptions written in Latin as well as in vernacular languages between the fourth and eleventh centuries, are still awaiting publication^{5,6}. Finally, even when adequate critical editions exist, as in the case of Pseudo-Apuleius' *Herbarius*⁷, few translations are available, making these works of little use to researchers who do not have a solid background in Greek or Latin.

The second problem is one of plant identification. How do we establish the identity of medicinal plants mentioned in texts that, as a rule, provide no iconography and only cursory botanical descriptions? Although the works of Theophrastus and Dioscorides provide a firm starting point and have made possible the compilation of important repertoires of classical phytonyms⁸, several problems remain. For example, how confident can we be that the same plant name used in the second century BC by the Italian Cato the Elder and in the fourth century AD by the North-African Pseudo-Apuleius indicate the same botanical species? Clearly, a detailed analysis should be carried out to confirm, case by case, the most plausible identifications.

The third problem, and possibly the most delicate, is one of historical perspective. We are used to thinking of pharmacological therapy as something exclusively rational and experimental. This may or may not be true for current therapy, but it is certainly false for premodern therapy. Like us, the Ancients believed in the value of experimental evidence. But they attributed to this deceptively simple concept meanings which, beside being widely different from our own, also changed considerably during the sixteen centuries that separate Hippocratic writers from, say, Hildegard of Bingen. Yet appreciating the various elements that made up premodern therapeutic choices — sometimes an assembly of empirics, philosophy and folklore — is essential to finding the way through the intricate pharmacopeias of the Ancients.

In conclusion, Holland's proposal opens up a promising avenue of research that may lead to the discovery of novel therapeutic agents (or rather, to the rediscovery of ancient ones). The problems outlined here underscore the importance, pointed out by Holland¹, of carrying out this research through a multidisciplinary effort involving specialists from such diverse fields as philology, botany, pharmacology and history of medicine.

Daniele Piomelli

Unité de Neurobiologie
et Pharmacologie de l'INSERM,
Paris, France

Antonino Pollio

Dipartimento di Biologia Vegetale,
Università Federico II,
Napoli, Italia

1. Holland, B. K. *Nature* **369**, 702 (1994).

2. Lloyd, G. R. *Science, Folklore and Ideology*, 112–217 (Cambridge University Press, 1983).

3. Marcellus Empiricus *De Medicamentis* (ed. Elmerich, G.) (Teubner, Lipsia, 1889).

4. Theodorus Priscianus *Euporiston* (ed. Rose, V.) (Teubner, Lipsia, 1894).

5. Sigerist, H. E. *Bull. Hist. Medicine* **2**, 26–52 (1934).

6. Stannard, J. *Bull. Hist. Medicine* **46**, 455–467 (1972).

7. Sigerist, H. E. & Howald, E. *Pseudo-Apulei Herbarius Corpus Medicorum Latinorum*, vol. IV (Leipzig and Berlin, 1927).

8. André J. *Les Noms des Plantes dans la Rome Antique* (Les Belles Lettres, Paris, 1981).

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only.

My first five years in science

Linus Pauling

Linus Pauling, a giant of modern chemistry, died on 19 August (see page 584 of last week's issue). What follows is an account, in his own words, of his first years as a research scientist.

SEVENTY-TWO years ago, in the spring of 1922, I was completing my fourth year in the Oregon Agricultural College. In June 1922 I received the degree of Bachelor of Science in chemical engineering. I had learned a little mathematics, physics and a moderate amount of engineering. In addition I had amassed a tremendous amount of information about the properties of substances, especially inorganic compounds, including minerals and alloys. I had become interested in the electron theory of valence in 1919, and I continued to hope that the empirical information about the properties of substances could eventually be encompassed in a theory of the structure of molecules.

I had also applied to several universities for work, and had accepted an appointment in the California Institute of Technology (Caltech). My introduction to research began in September 1922, when I arrived in Pasadena, and I consider that my education continued for 5 years, until 1927, when I returned from Europe.

During my three and a half years as a graduate student I had learned a great deal about modern physics, about mathematics, and about chemistry, especially chemical thermodynamics (which did not appeal to me) and statistical mechanics, which I liked very much and still like. Caltech was a remarkable place. The first PhD had been given in 1920, only 2 years before my arrival. The faculty contained many outstanding researchers in mathematics, physics and chemistry, including Roscoe Gilkey Dickinson (a pioneer X-ray crystallographer, the first recipient of a Caltech PhD), Harry Bateman, Paul Epstein, Robert A. Millikan, Arthur Amos Noyes and Richard C. Tolman. There was great interest in line spectroscopy, the old quantum theory, and other aspects of modern physics. In the chemistry department, led by Noyes, graduate students were expected to begin research immediately on their arrival in Pasadena, and in addition to study some advanced courses, many taught by Tolman. I especially enjoyed Tolman's course in statistical mechanics.

In addition, there was a weekly seminar in chemistry, a seminar in physics, and a seminar held at either the Caltech campus or the Mount Wilson Observatory Building in Pasadena by the physics and astronomy club. All these seminars involved the presentation of striking, newly published discoveries, such as the Compton

effect, the Stern-Gerlach experiment showing orientation of angular-momentum vectors in space, or the spin of the electron. My interest in X-ray crystallography has continued to the present. In addition, I attended courses of lectures by visiting physicists: Sommerfeld, Born (on matrix mechanics), Ehrenfest, Raman and Charles Galton Darwin. Although most of my early publications were based on experimental work, the determination of the structure of crystals, I soon began carrying out theoretical research.

I have had good luck several times during my life. One remarkable bit of good luck was in having gone to Caltech for my graduate studies. Years later, after I had become well acquainted with many outstanding universities, I realized that there was no place in the world in 1922 that would have prepared me in a better way for my career as a scientist.

After three and a half years in Pasadena I was sent to Europe as a research associate of Caltech, an arrangement made by Noyes. I had also applied for a John Simon Guggenheim Memorial fellowship, and shortly after my wife and I had arrived in Munich, I heard that the fellowship had been granted. We remained in Europe, one year in Munich with Sommerfeld and shorter stays in Copenhagen and Zurich, as well as visits to other centres of research in modern physics, for a year and a half. I had been given an appointment as an assistant professor at Caltech, and we returned to Pasadena in September 1927.

My education had continued during this period, especially during the year in Sommerfeld's Institute of Theoretical Physics. Just as we arrived, in April 1926, Schrödinger had begun publishing his papers on wave mechanics, and Sommerfeld immediately began giving lectures on them. Several other seminars were also devoted to newly published papers in wave mechanics and other aspects of quantum mechanics. I immediately began thinking about my main interest, which was not the hydrogen atom or hydrogen molecules, but rather the question of the structure and properties of many-electron atoms and ions and of more complex molecules. My paper was published in *Zeitschrift für Physik* for 1927. It involved an interesting combination of old quantum theory and wave mechanics, published by Gregor Wentzel earlier in 1926. The treatment involved the use of an atomic model developed in 1920 by Schrödinger, who

suggested that atoms and ions containing many electrons might be treated by an idealized model in which the inner shells of electrons were replaced by surface charges of electricity, with a single electron in an orbit penetrating these shells. Wentzel used Sommerfeld's old-quantum-theory method of quantizing conditionally periodic systems, with the replacement of the square of the orbital-angular-momentum quantum number by the product of the quantum number and the quantum number plus 1. Wentzel had evaluated the screening constants for X-ray doublets by this theoretical treatment, but his calculated values disagreed with the experimental ones. I noticed that an extension of Wentzel's method could be carried out which led to reasonably good agreement with experiment. So far as I know, this was the first application of quantum mechanics to atoms and ions containing many electrons.

I then continued to develop this method and applied it to the calculation of theoretical values of the electric polarizability, diamagnetic susceptibility, extension in space of many-electron atoms and ions, and also to the evaluation of a set of values of radii of ions for use in ionic crystals. These papers were published in 1927.

The fact that I was in Sommerfeld's institute just at the time that Schrödinger published his papers on wave mechanics provides a second example of my astonishing good luck. So far as I know, Sommerfeld's lectures on wave mechanics were the first on this subject anywhere in the world. It was somewhat by chance that I arrived in Munich in the spring of 1926; I had in fact thought seriously of going to Niels Bohr's institute in Copenhagen, but circumstances caused me to change my mind. I am confident that this bit of good luck was largely responsible for my having got a rather thorough grounding in wave mechanics within a few months of its discovery, and probably responsible for my later success in applying quantum mechanics to some previously puzzling questions in the field of chemistry, culminating in the publication of my book, *The Nature of the Chemical Bond*, in 1939.

There have been other times in my life when I have had good luck and some when I have been less fortunate; but I see no reason to complain. I look back on the 5 years beginning in September 1922 with much pleasure, and with great gratitude to my remarkable teachers. □

The return of cosmological creation

Without much notice, the notion that the Universe may have been through an infinite series of bouts of creation, and not just one, has been making headway in the literature.

SIR Fred Hoyle's autobiography *Home is where the Wind Blows*, has been well and enthusiastically reviewed, and for good reason; it is a stimulating read. Cosmologists will not be surprised to find that, among the often revealing gossip, there is a modest reiteration of Hoyle's distaste for the 'Big Bang' origin of the Universe, and his devotion to the notion of the continuous or at least repetitive creation of mass. But nowhere has it been remarked that, after a long fallow period, Hoyle appears again to be back in serious business, with a steady trickle of papers in the astronomy journals, the latest of them in this month's issue of *Astronomy and Astrophysics* (287, 729–739; 1994). What follows is an account of what he and his colleagues now have to say.

A convenient starting-point is the article by Arp *et al.* (*Nature* 357, 287; 1992), which was both a criticism of the Big Bang Universe and a tentative advocacy of the view that the appearance of matter in the Universe should be accommodated somehow in the laws of physics. That skeleton was covered with flesh just over a year ago, in an article in *The Astrophysical Journal* (410, 437–457; 1993) by Hoyle and his associates of long-standing, Geoffrey Burbidge from the University of California at San Diego and J. V. Narlikar, now director of the Inter-University Centre for Astronomy and Astrophysics at Pune, India.

No longer is the objective to support the steady-state cosmology of the 1950s, Hoyle's joint invention with Thomas Gold and Hermann Bondi, which had matter (presumably in the form of hydrogen atoms) trickling into the Universe at such a rate that its average density remained constant. Too much has happened since the 1950s for that to be a tenable proposition, notably the discovery of quasars and other active galactic nuclei, the discovery of the microwave background and the recognition that the true constituents of matter are quarks and leptons (electrons and the like).

The new theory is called, instead, a "quasi-steady state" cosmology. Matter does not trickle into the Universe, but instead appears only at places where its concentration is already high, in the nuclei of galaxies for example. What this means for the Universe as a whole is perhaps most easily gleaned from an article by the same three authors in the *Monthly Notices of the Royal Astronomical Society* (which now belies its title by appearing twice a month) in April

this year (267, 1007–1019; 1994).

In the large, the Universe is expanding, but there is superposed on it an oscillation of scale with a period of 40 billion years. The appearance of matter is concentrated at those epochs in the history of the Universe, stretches indefinitely backwards in time, when the scale of the Universe is near an oscillatory minimum, the most recent of which occurred 14 billion years ago, or not very different from the usual estimates of the age of the Big Bang.

Among other things, this has the consequence that the maximum redshift of objects (such as galaxies) making their appearance in the current oscillatory cycle will have a redshift parameter (called z) less than 4.86, but they are no longer the only objects in the Universe that are in principle observable. There is no reason why galaxies from earlier cycles of creation should not also be found; those left over from the immediately previous cycle will have redshifts less than 5.166, but will usually be three (astronomical) magnitudes fainter than objects from the current cycle, and thus will be more difficult to find. But Hoyle, Burbidge and Narlikar argue that radiogalaxies from earlier cycles should be more easily detectable at long wavelengths, which may help to explain the high frequency of very faint radiogalaxies now found.

In this light, the quasi-steady state cosmology may be thought to include the Big Bang as a special case. The Universe, while expanding overall, is one in which there are periodic outbursts of matter-creation the most recent of which was that 14 billion years ago. And does it not make sense to suppose that the direct evidence of that is the apparent outpouring of mass from the distant quasars, with redshift z in excess of 3.0 or more? In the prosecution of iconoclasm, it always helps if the defenders of a conventional view are able to retreat gracefully, by arguing that their own position has merely been absorbed in a more general framework.

Nevertheless it seems improbable that the flow of papers from Hoyle, Burbidge and Narlikar will quickly sweep away the Big Bang. Not that the authors have shirked the task, indispensable for iconoclasts, of attempting to show that the observations counted as evidence for the Big Bang are also compatible with their view. In this spirit they deal with the synthesis of elements such as helium, most of which, on any view, must be primordial and not (like heavier

elements) synthesized in stars. The microwave background arises, they say, from the scattering of radiation from earlier cycles of matter-creation; twenty are necessary to account for the present properties of the radiation field (which implies that the bulk of its photons began life 800 billion years ago).

And how is all this accommodated within the laws of physics? By a field, of course: a scalar field called C (for "creation") whose value is a function of space-time and which has the disconcerting (but not implausible) effect of exerting a negative pressure on matter. That idea was an essential ingredient of the original steady-state theory. The overall rate of creation is determined by the square of the time-derivative of C averaged over the whole Universe.

The particles associated with the field are bosons which, the argument goes, are converted at places where the gravitational field is very large into primordial matter particles, each with a 'Planck mass' of 10^{-5} grams, which are then converted into baryonic matter (quarks and so on) through the standard hierarchical relationship between the fundamental particles, as in the early phase of the Big Bang.

Whatever the critics may say, they will have to allow that it is a considerable *tour de force* on the part of Hoyle, Burbidge and Narlikar to have put so much flesh on an awkward skeleton without more than the most minimal institutional support, and on the basis of occasional weeks holed up together in some hotel.

But what will matter, in the end, is whether they can make some prediction about the Universe that will be arresting enough to divert the attention of others from more conventional trains of thought. This month's paper (in *Astronomy and Astrophysics*) says that there should be a predominance of blue galaxies, some with their light blue-shifted, among the faintest galaxies of all, which may be a beginning.

What Hoyle, Burbidge and Narlikar say is that they will next turn their attention to explosive creation (in their view, in quasars and the like) and see whether quasi-steady state cosmology has something to say about the distribution of galaxies as observed. However that comes out, they at least deserve credit for having pointed to one way in which the Big Bang, an event without a cause, might be brought within a wider framework.

John Maddox

The flight of the pterosaur

R. McNeill Alexander

How wide were pterosaur wings, and were they attached to the hind limbs (as in bats) as well as to the fore limbs? These are controversial questions, which if resolved would tell us a great deal about pterosaur flight performance and possible



FIG. 1 Fossilized flight gear — a specimen of *Sordes pilosus* from some 150 million years ago. (D. Unwin.)

running ability. On page 62 of this issue¹, Unwin and Bakhurina come up with some answers, in that they present evidence that the wing was attached to the hind limbs, right down to the ankles. Their interpretation of the species concerned, *Sordes pilosus*, incidentally implies that the hairy pterosaur was not hairy.

Early reconstructions of pterosaur wings were bat-like², but Padian^{3,4} studied the impressions of wing membranes found in a substantial number of fossils. These showed an acutely pointed wing tip, and a membrane that is narrow at least from the tip to the wrist joint. Padian argued that the whole wing must have been narrow and that its rear edge would have been attached to the trunk, leaving the hind legs free for fast bipedal running. He drew attention to fine fibres in the membrane, aligned like the primary feathers of birds, which he supposed to have stiffened the wing, preventing it from billowing upwards without any need to stretch it tight. His views have been widely, but not universally⁵, accepted.

The latest move in the controversy involves one of the most intriguing pterosaurs, the supposedly hairy *Sordes pilo-*

sus. Beautifully preserved specimens of this tern-sized species (see Fig. 1) have impressions of sparse wavy fibres on the inner part of the wing which have been interpreted as fur, apparently the most direct evidence of 'warm-bloodedness' in any Mesozoic reptile.

The important point in the present context, however, is that in *Sordes*, unlike other pterosaurs, impressions of the inner part of the wing membranes survive. The membranes apparently extended down the legs as far as the ankles and were continued as a membrane joining the two legs ventral to the tail. The outer parts of the wing membranes have close-packed, parallel fibres, as in other pterosaurs, but the inner parts have only sparser, less regularly arranged fibres (previously identified as hairs). The sparseness of the fibres in the inner part may explain why it has disappeared without trace in other pterosaur fossils.

Unwin and Bakhurina's reconstruction of the wings of *Sordes* is narrow near the tips and broad at the bases. This shape is quite different from any of the alternatives considered in a discussion of pterosaur flight performance by Hazlehurst and Rayner⁶, but its area cannot have been much different from their alternative (2) (see Fig. 2; in making this statement I have taken account both of the area of the membrane between the legs, and of a difference of hip and knee angles between the reconstructions). Unwin and Bakhurina interpret the outer (more fibrous) part of the membrane as relatively stiff, and the inner part as more flexible.

The aerodynamic performance of aircraft and flying animals depends very largely on two quantities calculated from wing dimensions, the wing loading and the aspect ratio. Wing loading is body weight divided by wing area. Aircraft with low wing loadings (relatively large wings) fly best at low speeds, whereas those with high wing loadings fly fast. Those with lower wing loadings are more manoeuvrable. Aspect ratio is wing span divided by mean chord (the width of the wing from front to rear). High aspect ratio (long narrow) wings give the best aerodynamic performance.

The wing dimensions of birds⁷ and bats⁸ are related to their flying habits in predictable ways. Fast fliers such as ducks have high wing loadings and slow fliers such as vultures have low ones. Sea birds such as albatrosses have high aspect ratios, whereas birds that fly in cluttered environments such as woods (where long wings would be apt to hit branches) have low aspect ratios.

Where does this leave us with respect to pterosaur flight? Take the estimated masses and wing spans of pterosaurs from ref. 6 and assume that the wing areas were as given for model (2) wings. Then the wing loadings of small pterosaurs were lower than those of most birds of similar mass, and about typical for similar-sized bats. The aspect ratios were fairly high, in comparison both with birds and with bats. This suggests that the pterosaurs were slow, highly manoeuvrable fliers, and were best suited (because of their long wings) for flying in uncluttered habitats. More specifically, the pterosaurs *Pterodactylus micronyx* (estimated mass 27 g), *Scaphognathus crassirostris* (43 g) and *P. antiquus* (96 g) were closely similar both in wing loading and in aspect ratio

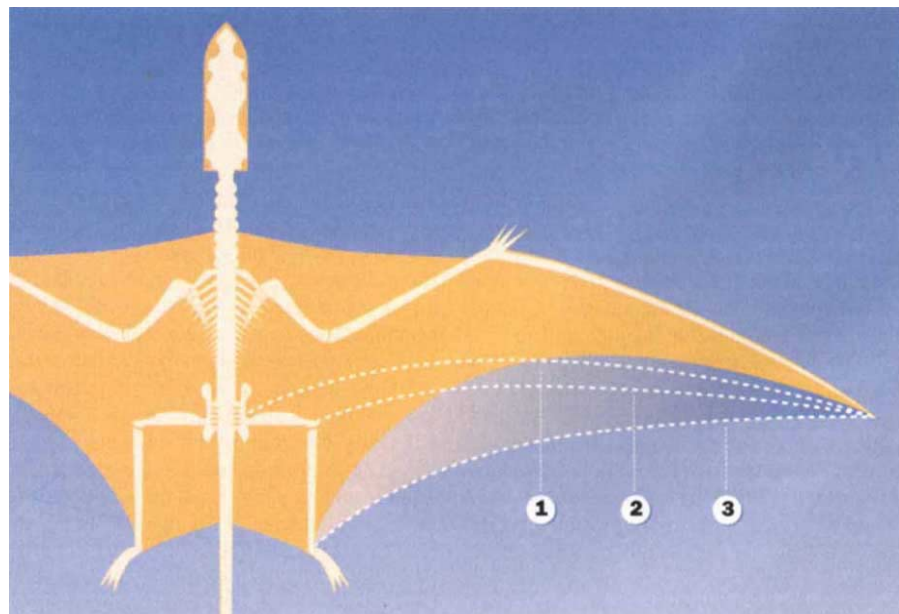


FIG. 2 Possible pterosaur wing patterns. Hazlehurst and Rayner's⁶ three alternative wing restorations superimposed on the reconstruction of Unwin and Bakhurina¹.

to swallows (*Hirundo*)⁹, fish-eating bats (*Noctilio*)⁸ and terns (*Sterna*)⁹, respectively of equal mass. These similarities seem understandable given that these pterosaurs are found in marine or lagoon deposits, and in several cases have remains of fish in the gut⁶. The nearest modern equivalents to *Pteranodon* (17 kg) and the other large pterosaurs seem to be frigate birds (*Fregata*), which soar in convection currents over the sea, hunt-

ing for fish and squid¹⁰.

If the wing membranes were indeed attached to the ankles, it seems more likely that pterosaurs crawled like bats, than that they were the fast runners of Padian's reconstruction. Which seems a pity. □

R. McNeill Alexander is in the Department of Pure and Applied Biology, University of Leeds, Leeds LS2 9JT, UK.

made observations of C⁰ in the $z = 1.776$ absorption-line system towards the quasar 1331+170. In a 13-hour net exposure with the massive high-resolution spectrograph, they not only detect the excited C⁰ absorption, but also resolve it into two velocity components indicative of two clouds within the absorbing system. The derived C⁰ excitation temperatures of 10.4 ± 0.5 K and 7.4 ± 0.8 K for the two clouds compare with the predicted CMBR temperature of 7.58 K at $z = 1.776$. The fact that the best previous C⁰ measurement was an upper limit of 16 K in the same 1331 + 170 system⁸ illustrates the potential for opening new scientific frontiers with the Keck telescope and others in the 10-m class.

It is very interesting that the C⁰ excitation in one of the clouds towards 1331+170 is identical within the measurement errors to the expected CMBR temperature at that redshift. As the presence of local excitation such as that due to hydrogen collisions can only serve to raise the C⁰ excitation above that of the CMBR contribution, this result indicates that local effects must be very small in at least one of the 1331+170 clouds if the standard Big Bang model is valid. Unfortunately, it is hard to pin down the amount of local C⁰ excitation in these clouds through other observations. Therefore, despite its agreement with previous expectations, Songaila and co-workers' measurement of 7.4 ± 0.8 K must be strictly considered an upper limit on the CMBR temperature at $z = 1.776$.

Nonetheless, the results are especially important in demonstrating that the local C⁰ excitation in at least some quasar absorbers is low enough to hope for a sensitive test of the standard Big Bang in higher-redshift quasar systems where the CMBR temperature should be even greater. If a set of such measurements consistently yields C⁰ excitation temperatures at or above the expected CMBR value, it would be another heavy stroke in favour of the Big Bang. On the other hand, the discovery of a system whose C⁰ excitation was significantly lower than the predicted CMBR temperature would pose a serious problem for cosmology. □

David M. Meyer is in the Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60208, USA.

1. Songaila, A. *et al.* *Nature* **371**, 43–45 (1994).
2. Smoot, G. F. *et al.* *Astrophys. J.* **396**, L1–L6 (1992).
3. Mather, J. C. *et al.* *Astrophys. J.* **420**, 439–444 (1994).
4. Thaddeus, P. *A. Rev. Astr. Astrophys.* **10**, 305–332 (1972).
5. Roth, K. C., Meyer, D. M. & Hawkins, I. *Astrophys. J.* **413**, L67–L71 (1993).
6. Bahcall, J. N. & Wolf, R. A. *Astrophys. J.* **152**, 701–729 (1968).
7. de Boer, K. S. & Morton, D. C. *Astr. Astrophys.* **37**, 305–311 (1974).
8. Meyer, D. M., Black, J. H., Chaffee, F. H., Foltz, C. B. & York, D. G. *Astrophys. J.* **308**, L37–L41 (1986).

1. Unwin, D. M. & Bakurina, N. N. *Nature* **371**, 62–64 (1994).
2. Wellnhofer, P. *The Illustrated Encyclopedia of Pterosaurs* (Salamander, London, 1991).
3. Padian, K. *Paleobiology* **9**, 218–239 (1983).
4. Padian, K. & Rayner, J. M. V. *Am. J. Sci.* **293A**, 91–166 (1993).
5. Pennycook, C. J. *Biol. Rev.* **63**, 299–311 (1988).

6. Hazlehurst, G. A. & Rayner, J. M. V. *Paleobiology* **18**, 447–463 (1992).
7. Rayner, J. M. V. *Curr. Ornithol.* **5**, 1–65 (1988).
8. Norberg, U. M. & Rayner, J. M. V. *Phil. Trans. R. Soc.* **B316**, 335–427 (1987).
9. Brooke, M. & Birkhead, T. *The Cambridge Encyclopedia of Ornithology* (Cambridge Univ. Press, 1991).
10. Alexander, R. McN. *Dynamics of Dinosaurs and Other Extinct Giants* (Columbia Univ. Press, New York, 1988).

COSMOLOGY

A distant space thermometer

David M. Meyer

As the redshifted fossil of the primeval fireball, the cosmic microwave background radiation (CMBR) is widely regarded as the best evidence that the Universe really did start with a Big Bang. A fundamental prediction of the standard Big Bang is that the temperature of this black-body radiation was higher in the past. In the first serious test of this prediction, on page 43 of this issue¹ Songaila *et al.* report a measurement of the CMBR temperature at large redshift that is strikingly consistent with the Big Bang theory.

The past few years have been exciting times for CMBR researchers. In particular, the Cosmic Background Explorer Satellite (COBE) has yielded data of unprecedented sensitivity on the isotropy² and spectrum³ of the CMBR. The latter results have convincingly confirmed the predicted black-body character of the spectrum and precisely determined the local present-day radiation temperature to be 2.726 ± 0.010 kelvin. However, the Big Bang prediction that the CMBR temperature increases linearly with redshift (z) by a factor of $(1+z)$ cannot be tested by COBE or any other direct measurement of the CMBR, as it requires one to measure the CMBR temperature in the past.

Fortunately, indirect measurements of the CMBR are possible by observing the excitation of certain molecules and atoms in interstellar and extragalactic gas clouds. Indeed, some of the most sensitive measurements of the CMBR temperature have been provided by optical absorption-line observations of the rotational excitation of interstellar cyanogen (CN) molecules towards nearby bright stars⁴. The most recent CN results⁵ yield a CMBR

temperature of $2.729^{+0.023}_{-0.031}$ K that is completely consistent with the COBE value. As the CN molecules provide a measurement of the CMBR at the locations of their host interstellar clouds several hundred light-years distant, they sample the CMBR as it was several hundred years ago. In itself, this time base is far too short to test the predicted thermal evolution of the CMBR, but it illustrates an idea that could be applied to the high-redshift clouds seen in absorption in the spectra of distant quasars.

Although CN has not been detected in any quasar absorption-line system to date, it has long been recognized⁶ that several atomic constituents of these clouds could serve as distant probes of the CMBR. With an energy separation corresponding to a wavelength of 0.61 mm, the fine-structure splitting of atomic carbon should be particularly sensitive to the CMBR in high-redshift ($z > 1.5$) quasar absorbers. As the flux of a 2.73 K black body is weak at 0.61 mm, collisions with hydrogen atoms are much more important than the CMBR in populating the excited fine-structure levels of atomic carbon (C⁰) in galactic interstellar clouds⁷. But the carbon atoms should be appreciably excited by the CMBR at large redshifts where the temperature is presumably higher and the peak of its flux is nearer to 0.61 mm. The difficulty in testing this hypothesis is that the neutral carbon absorption is weak and that suitably distant quasars are at least 10,000 times fainter than the faintest stars used as background sources in the interstellar CN studies.

Songaila *et al.* have succeeded in bringing the tremendous photon-gathering power of the Keck 10-m telescope to bear on this long-standing problem. They have

Genocide in tribal society

Shepard Krech III

THE signatories of The Unanimous Declaration of the Thirteen United States of America, dated 4 July 1776, had a very large number of complaints against the king of Great Britain. Among them was this: "He . . . has endeavoured to bring on the inhabitants of our frontiers, the merciless Indian Savages, whose known rule of warfare is an undistinguished destruction of all ages, sexes and conditions". Taking that statement on trust — a risky business at best — a big question still remains. Was this form of warfare practised before Europeans started to settle in North America, or was it introduced by them?

In recent years, anthropologists interested in explaining war have actively investigated this very problem. As Douglas Bamforth points out in the anthropological journal *Man*¹, a consensus has developed that warfare waged in state-level societies contrasts fundamentally with war conducted in tribal societies. In the first, war is especially brutal and usually undertaken to exterminate an enemy who speaks another language, possesses a separate culture, and can readily be dehumanized. This genocidal form of warfare is said to be rare in tribal societies, where war is much tamer and its objectives are not to inflict casualties but to gain honours and other ends — rare, that is, until those societies come into contact with nation-states and at the same time acquire the means to wage wars of annihilation^{2,3}. Guns often head the list of such means. Thus Chagnon on the Yanomamo of Amazonia: "The possession of the gun caused war where none had previously existed" (ref. 4). Guns, of course, did not appear in the tribal world until nation-states supplied them⁵.

Against this are the views that, regardless of the presence or absence of the West, especially virulent personalities predispose men in certain tribal societies to war; that explanations should be sought in the clear Darwinian advantage, in numbers of offspring, which go to the victors in war; that revenge is motive enough for blood feuds of long standing and especial ferocity; or that there are always compelling economic and ecological reasons for war⁶. But these opinions are on the defensive, because evidence lending support to a particular theory is either unconvincing or postdates the arrival of state-level societies.

Bamforth is the latest to take issue with

the prevailing consensus. He is an archaeologist and he brings together historical and archaeological evidence from one part of the North American Great Plains, from the period between the twelfth and nineteenth centuries, to discuss war both before and after the coming of Europeans. It is the archaeological data from early fourteenth century sites that

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Utes at Los Pinos, Colorado, circa 1880. Note that guns now replace the traditional bows and arrows. The Ute are a North American Indian people of Utah, Colorado and New Mexico who are related to the Aztecs. From the Hayden Album of North American Indians.

are especially important, because they indicate the occurrence of a decidedly nasty event long before Europeans arrived.

The archaeological sites are located along the Missouri River and were inhabited by the sedentary agricultural ancestors of people such as the Siouian-speaking Mandan and Hidatsa, and the Caddoan-speaking Arikara and Pawnee. The Caddoans moved into the region near the turn of the fourteenth century, at which time it seems that their villages and the Siouian villages became heavily fortified and placed strategically with a mind to defence. At one site in South Dakota, named Crow Creek, the consequences of inadequate fortifications or defence are clear: almost 500 people of both sexes and all ages were killed and their village was burned. The inhabitants were mutilated before or after death: half of the skulls were fractured and over 90 per cent scalped, bodies were decapitated, and extremities were cut from limbs and tongues from throats⁷⁻⁹. If they were Siouian, the Indians who laid waste to the Caddoans of Crow Creek possessed a different

language and culture to those of their victims. They waged war we would unhesitatingly call genocidal if we knew that it involved the deliberate extermination of a people; the Crow Creek villagers were a 'people', and with 500 dead they may well have been exterminated. But we cannot know if the action was deliberate.

Bamforth's analysis provides a corrective to the idea that war on the Great Plains was always some kind of grand diversion, as deadly as a vigorous game of tag (see also ref. 10). More broadly, it runs against the related dogma that before

Royal Anthropological Institute

Columbus arrived in North America warfare was universally ritualized or restrained. This idea became newly popular as 1992, the year of the Columbus quincentenary, approached. It has entered public culture through those who simultaneously wish Eden or the Islands of the Blessed upon North America, and evils upon Europe. Crow Creek puts the lie to the statement that "there is no very good proof" that warfare before the coming of Europeans was "particularly fierce" (refs 11, 12).

Bamforth suggests that Crow Creek was attacked over access to productive resources. Good land for crops was at a premium in this part of the plains, and both the residents of Crow Creek and their enemies were maize-dependent horticulturalists with expanding populations. Bone analysis indicates that the people of Crow Creek suffered from episodic malnutrition, iron deficiency and metabolic defects. Perhaps the victors also were malnourished and scarred osteologically, but there is no firm evidence. Maybe they acted to correct the conditions, as Bamforth suggests. But we do not know and may never know for certain why the residents of Crow Creek were massacred.

Moreover, despite the temptation to draw broader inferences from Crow

1. Bamforth, D. H. *Man* **29**, 95–115 (1994).
2. Blick, J. P. *Man* **23**, 654–670 (1988).
3. Ferguson, R. B. & Whitehead, N. (eds) *War in the Tribal Zone* (School of American Research, Santa Fe, 1992).
4. Chagnon, N. *The Yanomamo* 3rd edn, 203 (Holt, Rinehart & Winston, 1983).
5. Ferguson, R. B. (ed.) *Warfare, Culture, and Environment* (Academic, New York, 1984).
6. Ferguson, R. B. in *Warfare, Culture, and Environment* (ed. Ferguson, R. B.) 1–81 (Academic, New York, 1984).
7. Kivett, M. & Jensen, R. *Archaeological Investigations at the Crow Creek Site 38BF11* (Nebraska State Historical Soc. Publ. in Anthropology 7, 1976).
8. Zimmerman, L. et al. *The Crow Creek Site (38BF11) Massacre: A Preliminary Report* (US Army Corps of Engineers, Omaha District, 1981).
9. Willey, P. *Prehistoric Warfare on the Great Plains* (Garland, New York, 1990).
10. Biolsi, T. in *Warfare, Culture, and Environment* (ed. Ferguson, R. B.) 141–168 (Academic, New York, 1984).
11. Sale, K. *The Conquest of Paradise* 318 (Plume, New York, 1990).
12. Holm, T. in *The State of Native America* (ed. Jaimes, M. A.) 355–356 (South End Press, Boston, 1992).

Creek and the neighbouring palisaded sites, one cannot guess at how frequently aboriginal war was waged here or anywhere else, as Bamforth himself realizes. Because of the nature of archaeological preservation, especially in tropical America, the larger picture is not likely to change soon. Nor does Crow Creek negate the evidence for destructiveness of war waged by native Americans dislocated, pressured or stimulated by the presence of state-organized systems, in-

cluding those from Europe. Nonetheless, this case does show that neither guns nor the presence of a state (European or other) were necessary for a form of aboriginal warfare whose aim, it may reasonably be inferred, corresponded with the result at Crow Creek — the extirpation of an entire community. □

Shepard Krech III is in the Department of Anthropology, Brown University, Providence, Rhode Island 02912, USA.

DEVELOPMENTAL NEUROBIOLOGY

Attractive guides for axons

Dennis D. M. O'Leary

In two ground-breaking papers in *Cell*, both from Marc Tessier-Lavigne's group, Serafini *et al.*¹ and Kennedy *et al.*² report the isolation and cloning of diffusible chemotropic factors endogenous to the vertebrate nervous system. The factors, named netrin-1 and netrin-2 ('netr' derives from the Sanskrit for 'one who guides'), promote and reorient the growth of spinal commissural axons *in vitro*. These properties mimic those described for a diffusible activity released from the floor plate of the spinal cord in embryonic chickens and rodents^{3,4}, suggesting that the netrins play a key role in guiding the circumferential growth of commissural axons from their origin in dorsal spinal cord to the ventral midline.

The issue of axon guidance is a fundamental one in developmental neurobiology. The idea that a diffusible activity released from a distant source can direct growing axons was first put forward over a century ago by the great Spanish neurobiologist, Ramon y Cajal⁵; the idea lay dormant until it was resurrected in part by Gunderson and Barrett⁶, who demonstrated that sensory axons grown *in vitro* will turn towards a source of nerve growth factor. Although nerve growth factor does not act *in vivo* to attract developing sensory axons to their peripheral targets⁷, the experiment reopened the investigation of chemotropism. In pivotal studies^{8,9}, involving co-culture in collagen gels of the trigeminal ganglion with its peripheral target, the maxillary epithelium, it was shown that the epithelium releases a diffusible factor which attracts growing trigeminal axons. The role of chemoattraction in the development of neural connections in the central nervous system was later indicated by the demonstration that the floor plate releases a diffusible activity that influences the growth of spinal commissural axons^{3,4}, and that the basilar pons, a brainstem structure, releases a diffusible activity that promotes its innervation by corticospinal axons^{10,11}.

It was not feasible to isolate the chemo-

tropic activity directly from the floor plate, given its small size. After screening several neural tissues and cell lines for a floor-plate-like activity, two proteins with relative molecular masses of 78K and 75K (later named netrin-1 and -2) were purified from the membrane fraction of embryonic chick brain on the basis of their ability to promote neurite outgrowth from explants of dorsal spinal cord¹. Partial amino-acid sequences of the purified proteins were used to design oligonucleotide primers for isolating complementary DNA fragments by polymerase chain reaction. Subsequent screening of libraries of embryonic chick brain and spinal cord yielded cDNAs encoding the two proteins. Both molecules have an amino-terminal signal sequence consistent with a secreted protein¹.

Several lines of evidence indicate that netrin-1 is the floor-plate chemoattractant². First, floor-plate cells express netrin-1 RNA. Second, netrin-1 mimics the activity released by floor plate. Floor plate had already been shown to promote at a distance the outgrowth of commissural axons from explants of dorsolateral spinal cord³. In addition, pieces of floor plate juxtaposed to spinal cord explants induced the turning of commissural axons within the spinal explant⁴. Both of these effects of the floor plate on commissural axons are replicated by recombinant netrin-1 secreted by transfected COS

cells. In the collagen gel assay, the action of netrin-2 is similar to that of netrin-1, but their expression patterns within the spinal cord are complementary: *netrin-1* expression is localized to the floor plate, whereas *netrin-2* is not expressed in the floor plate but at a low level over the ventral two-thirds of the spinal cord — a distribution coincident with the path taken by commissural axons to the midline. It will be interesting to sort out the relative contributions of the two netrins in directing the growth of spinal commissural axons, as well as of other axonal populations, in view of the fact that each gene is also expressed in restricted domains outside the spinal cord².

Drosophila and *Caenorhabditis elegans* genetics have contributed much to identifying genes involved in axon guidance, with vertebrate studies until now largely defining mechanisms at an anatomical and cellular level. A number of molecular activities, both diffusible and membrane-bound, have been discovered in a variety of bioassays using vertebrate tissues, but identifying the molecules and genes responsible for these activities has proved an arduous task. The first of these activities to be cloned was collapsin¹², a secreted protein purified from chick brain membrane extracts on the basis of an *in vitro* assay of growth cone collapse. Although a role for collapsin *in vivo* has yet to be defined, it may steer axons by inhibiting or repulsing them¹³.

The gene encoding collapsin was found to share homology with a previously cloned *Drosophila* gene, *fasciclin IV*, which encodes a transmembrane protein whose distribution suggests that it may have a function in axon guidance. The homology with collapsin suggested an inhibitory role for Fasciclin IV, consistent with axonal growth cone behaviours relative to Fasciclin IV domains in the limb bud of normal grasshoppers and following treatment with blocking antibodies against Fasciclin IV (refs 14, 15). Since then, other genes homologous to *collapsin* and *fasciclin IV* have been identified, which together form a family that encode transmembrane and secreted molecules known as semaphorins¹⁶.

This conservation between vertebrates

- Serafini, T. *et al.* *Cell* **78**, 409–424 (1994).
- Kennedy, T. E., Serafini, T., de la Torre, J. R. & Tessier-Lavigne, M. *Cell* **78**, 425–436 (1994).
- Tessier-Lavigne, M., Placzek, M., Lumsden, A. G. S., Dodd, J. & Jessell, T. M. *Nature* **336**, 775–777 (1988).
- Placzek, M., Tessier-Lavigne, M., Jessell, T. M. & Dodd, J. *Development* **110**, 1930–1942 (1990).
- Ramon y Cajal, S. *Le Cellule* **9**, 119–258 (1892).
- Gundersen, R. W. & Barrett, J. N. *Science* **206**, 1079–1080 (1979).
- Davies, A. M. *et al.* *Nature* **326**, 353–363 (1987).
- Lumsden, A. G. S. & Davies, A. M. *Nature* **306**, 786–788 (1983).
- Lumsden, A. G. S. & Davies, A. M. *Nature* **323**, 538–539 (1986).
- Heffner, C. D., Lumsden, A. G. S. & O'Leary, D. D. M. *Science* **247**, 217–220 (1990).
- Sato, M., Lopez-Mascaraque, L., Heffner, C. D. & O'Leary, D. D. M. *Neuron* (in the press).
- Luo, Y., Raible, D. & Raper, J. A. *Cell* **75**, 217–227 (1993).
- Fan, J. & Raper, J. A. *Soc. Neurosci. Abstr.* **19**, 236 (1993).
- O'Connor, T. P., Duerr, J. S. & Bentley, D. J. *Neurosci.* **10**, 3935–3946 (1990).
- Kolodkin, A. L. *et al.* *Neuron* **9**, 831–845 (1992).
- Kolodkin, A. L., Matthes, D. J. & Goodman, C. S. *Cell* **75**, 1389–1399 (1993).
- Hedgecock, E. M., Culotti, J. G. & Halls, D. H. *Neuron* **4**, 61–85 (1990).
- Ishii, N., Wadsworth, W. G., Stern, B. D., Culicotti, J. G. & Hedgecock, E. M. *Neuron* **9**, 873–881 (1992).
- Colamarino, S. A. & Tessier-Lavigne, M. *Soc. Neurosci. Abstr.* (in the press).
- Pini, A. *Science* **261**, 95–98 (1993).
- Fitzgerald, M., Kwiat, G. C., Middleton, J. & Pini, A. *Development* **117**, 1377–1384 (1993).
- Mitchell, K. *et al.* *Soc. Neurosci. Abstr.* (in the press).

and invertebrates, not only in gene sequence but also in potential gene function, is highlighted by the homology of the netrins to *unc-6*, a gene that has been cloned in the nematode and encodes a laminin-related protein^{17,18}. Mutants for *unc-6* show guidance deficits in the circumferential growth of pioneer axons and migrations of mesodermal cells. This uncanny similarity suggests that UNC-6 protein may serve as a chemoattractant in nematode development in much the same way as netrins may control the development of circumferential commissural projections in vertebrate spinal cord.

An intriguing twist in the netrin story has developed from the discovery that genetically separable mutations of *unc-6*, which affect different domains of the protein, result in selective deficits in the nematode: for example, disrupted migration of mesodermal cells but not axons, or of ventral but not dorsal migrations or vice versa¹⁷. This latter finding in particular suggested to Tessier-Lavigne and colleagues that the netrins may have biological activities apart from chemoattraction. One possibility is that netrins may inhibit

or repulse the growth of certain axons. *In vivo*, trochlear motor neurons are located near the midline in the ventral part of the caudal midbrain, and their axons take a dorsolateral pathway to exit from the dorsal surface of the midbrain, forming the trochlear nerve. The floor plate, which expresses *netrin-1* along its entire length at the ventral midline throughout the spinal cord and into the hindbrain and midbrain², may influence the dorsolateral trajectory of trochlear axons by inhibiting or repelling their ventral growth. Indeed, preliminary findings indicate that both the floor plate¹⁹ and netrin-1-secreting COS cells (S. A. Colamarino and M. Tessier-Lavigne, personal communication) implanted into collagen gels inhibit neurite outgrowth from trochlear motor neurons. Thus the netrins may influence the directional growth of distinct axonal populations by two different mechanisms, chemoattraction and chemorepulsion. Only in the past year has chemorepulsion become a serious player in axon guidance, now that a diffusible activity released by the septum has been shown to repulse olfactory bulb axons in collagen gels; this

activity may act *in vivo* to steer olfactory axons away from the midline and form the laterally positioned olfactory tract^{20,21}.

The isolation and cloning of the netrins will have a considerable impact on developmental neurobiology, especially the issue of axon guidance. We await more on both general and specific mechanisms of axon guidance from analysis of netrin function and modes of action, identification of netrin receptors, and the search for other members of the *netrin/unc-6* family. Towards this end, a homologue of *netrin-1*, named *D-netrin*, has now been cloned in *Drosophila*; like its vertebrate counterpart, *D-netrin* is expressed in midline cells at a stage appropriate to influence commissure development²². Genetic analysis of *D-netrin* function in *Drosophila* should provide further insight into the role of netrins in axonal guidance, which is more than likely to be applicable to both invertebrates and vertebrates. □

Dennis D. M. O'Leary is in the Molecular Neurobiology Laboratory, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.

COMET SHOEMAKER-LEVY 9

A view from the far side

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

NASA JPL

BARELY a month has passed since the world was witness to perhaps the most dramatic astronomical event of the century. Over the course of a week, numerous fragments of comet Shoemaker-Levy 9 ploughed into Jupiter, leaving in their wake colossal plumes of dust and gas, and a band of dark scars in the giant planet's atmosphere. Although spectacular in their own right, Earth-bound observations were limited to the aftermath of the collisions and little is yet known about the impacts themselves.

The images above show our first glimpse of an actual collision in progress, as recorded by the Galileo spacecraft. Galileo is currently *en route* to Jupiter, scheduled to arrive in late 1995. At the time of the impacts the spacecraft was still some two billion kilometres from its destination (no doubt astronomers would have preferred it to have been closer), but it was nevertheless ideally placed to observe the collisions, on the rear side of Jupiter.

These first four images, taken 2½ seconds apart, show a bright flash of light signalling the beginning of the end of the

final piece of the comet, fragment W. This event took place some seven minutes before the accompanying impact plume emerged over Jupiter's limb and came into view from the Earth. Also visible on the illuminated half of Jupiter are several dark spots, the remnants of previous impacts.

Although flashes of light accompanying the impacts were widely expected, they do not appear to have been as dramatic as many had predicted. Indeed, none was detected from Earth. That Galileo has observed such flashes is, therefore, encouraging. But the data thus far returned raise several questions. It is not, for example, clear whether we have observed a fireball resulting from the catastrophic disruption of the fragment, or simply the luminous trail of the fragment as it plunged, meteor-like, into the jovian atmosphere.

As just a tiny fraction of the images recorded by Galileo have been returned to date, we can only hope that a more coherent picture of the events will emerge when more of the data are in hand.

Karl Ziemelis

Over the top to SUSY?

A. D. Martin

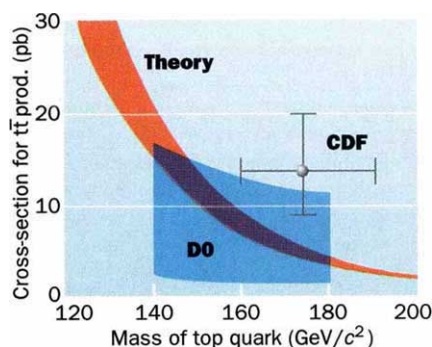
"If we had observed either twice or half the number of events then our 150-page paper¹ would have been much shorter" — thus H. Jensen (Fermilab) summed up the evidence for the top quark at the biennial high-energy physics conference, held this year in Glasgow*. The evidence for this, the elusive sixth quark whose detection has been anticipated ever since the discovery of the bottom quark in 1977, comes from the proton-antiproton Collider Detector at Fermilab (CDF) near Chicago. It was highly publicized in April (see the discussion in News and Views²), but the warts-and-all CDF presentation at the conference allowed delegates to evaluate and question the details. Certainly, at least twice the number of events are required to be certain that top has been discovered — half the events and there would be no evidence, but only a lower bound on the mass of the top quark. The second experiment at Fermilab (known as D0) has so far collected only just over half the data obtained by CDF; D0 does have a small excess of events over background, but as yet insufficient evidence to claim the identification of top (P. Grannis, SUNY, Stonybrook).

A top quark, t , decays into a bottom quark, b , and a W boson. In a proton-antiproton collision it is produced with its antiquark $\bar{t}$ and so a top event contains two W bosons and a $b\bar{b}$ pair which then have to be identified from their decays. From about 12,000 identified W boson events, the CDF experiment finds 15 top candidates (based on 12 events) of which 6 are expected to arise from other 'background' processes. The D0 experiment observes 7 possible candidates to be compared to 3.2 expected 'background' events. The compatibility of the results of the two experiments is seen in the figure, which shows the cross-section for $t\bar{t}$ production if different masses of the top quark are used. CDF has evidence for a top quark of mass $m_t = 174 \pm 10$ (statistical) $^{+13}_{-12}$ (systematic) GeV/c^2 , albeit with a cross-section that is more than twice the predicted value. If the small excess of D0 events is interpreted as top, then their cross-section lies within the darker blue band. The two experiments are again taking data and more results are eagerly awaited.

Why is it exciting to detect a quark that almost everyone expected to be there? The reason is that knowledge of its mass helps to answer the fundamental question of how the mass of elementary particles is generated. In the Standard Model the remaining undetected particle — the

Higgs boson — is responsible for generating mass.

The observed mass of the top can be compared with the indirect determination of its mass from the small contributions it makes by being 'exchanged' in a wide variety of precisely measured electroweak processes³. D. Schaile (CERN) reported a mass $m_t = 178 \pm 11^{+18}_{-19} \text{ GeV}/c^2$ from a global analysis of the latest measurements made, in particular, in the Z boson pro-



Evidence for a top quark of mass $m_t \approx 174 \text{ GeV}/c^2$ (at more than twice the predicted production rate) obtained by the CDF experiment. The darker blue band shows the expectations of the D0 experiment if their small excess of events were to be attributed to $t\bar{t}$ production.

duction and decay experiments at the Large Electron-Positron collider at CERN, Geneva, and also at the Stanford Linear Accelerator. Contributions from 'exchanged' or 'virtual' Higgs bosons also affect the electroweak processes but to a much smaller extent than those coming from the top quark. Nevertheless the error $^{+18}_{-19} \text{ GeV}/c^2$ on m_t arises from allowing the mass of the Higgs boson to vary from 1,000 down to 60 GeV/c^2 .

The quantum chromodynamic (QCD) sector of the Standard Model is also standing up to much more detailed scrutiny. For example, the many independent determinations of the QCD coupling are all consistent with a world average of $\alpha_s = 0.117 \pm 0.005$ (B. Webber, Cavendish Laboratory, Univ. Cambridge). There are also many new measurements from the HERA electron-proton collider in Hamburg (J. Feltesse, Saclay), which is now running with significantly improved interaction rate. The dramatic rise in the number of energetic (or so-called 'deep-inelastic') scatters as the electron probes the proton ever more finely, is now firmly established and allows QCD to be explored in an entirely new domain. Of particular interest are the 'diffractive' events seen in about 10% of the deep inelastic scatters. These may cast light on

the nonperturbative properties of the gluon, as could the new gluon-rich hadronic states which appear to exist in addition to those predicted by the quark model (C. Amsler, CERN).

The overriding impression from the conference was of the incredible all-round success of the Standard Model — a model that we nonetheless believe is incomplete. The most favoured, and theoretically elegant, extension comes in the form of supersymmetric Grand Unified Theories. Indeed the precision measurement of the weak mixing angle is in detailed agreement with the prediction of the simplest such extension, the supersymmetric-SU(5) model (S. Dimopoulos, CERN). Supersymmetry, or in brief SUSY, demands that each particle has a partner called a sparticle. There is, as yet, no evidence of these massive sparticles, and discovery of them would be a dramatic event indeed. It is a task ideally suited to the Large Hadron Collider which it is proposed to build at CERN.

The Standard Model strongly inhibits transitions between quarks of the same charge; that is, between down (d), strange (s) and bottom (b) quarks or between up (u), charm (c) and top (t). Rare decays based on such transitions are therefore an ideal hunting ground for new physics — undetected particles can influence the decay rate through their exchange contributions and hence be identified indirectly (just like the top quark). An observation made by the CLEO collaboration at the Cornell e^+e^- collider of the rate of a particularly important rare decay — the inclusive $b \rightarrow s\gamma$ decay — is of especial note (E. Thorndike, Univ. Rochester). The observed rate turns out to be entirely consistent with the Standard Model prediction. It therefore puts limits on the possible existence of charged Higgs bosons (whose presence would be an indicator of supersymmetry).

What of insights from astrophysics and cosmology? Three new independent measurements of the solar neutrino flux were presented at the meeting, each with errors of about 15%. There is general agreement that the observed flux is about 60% of that expected from the conventional solar model. If the solar model is accepted (and some question it), then the most obvious explanation is oscillations between the different types of neutrino, which may imply that neutrinos have mass. Another possible hint of neutrino oscillations comes from the new precision analysis of the cosmic ray data collected by the Kamiokande water Cerenkov detector in Japan, which finds only half the expected number of muon neutrinos.

Both the lightest sparticle (which is absolutely stable in many supersymmetric theories) and massive neutrinos could be abundant in deep space and make up the 'dark matter' that astronomers know ex-

*27th International Conference on High Energy Physics, Glasgow, 20–27 July 1994.

ists, because of the effects of its gravity, but that they cannot see. Hence the importance of the detection by gravitational (micro-)lensing of massive compact halo objects (MACHOs), which contribute to the baryonic dark matter, in our Galaxy (ref. 4; E. Kolb, Fermilab). Whether this leaves room for the existence of non-baryonic dark matter in the halo of our galaxy remains controversial⁵. There is however increasing evidence from the dynamics of galaxy clusters on large scales for large amounts of dark matter outside our own Galaxy, which could well be large enough to close the Universe. It would need to be non-baryonic (such as massive neutrinos and the lightest supersymmetric particle), because Big Bang nucleosynthesis implies that baryonic matter can contribute at most about 0.1 of the critical density.

Overall, there was a feeling of significant progress, that particle physics is entering a new era of international partnership to investigate the truly fundamental question of what lies beyond the Standard Model. Many believe that the recent experimental results give strong hints that SUSY will be embedded in the answer. □

A. D. Martin is in the Department of Physics, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, UK.

1. Abe, F. *et al.* (CDF collaboration) preprint FERMILAB-PUB-94/097-E, April 1994. To be published in *Phys. Rev. D* **50**, 1 Sept. 1994.
2. Close, F. & Maddox, J. *Nature News & Views* **368**, 805 (1994).
3. Ellis, J. *Nature News & Views* **370**, 101–102 (1994).
4. Hogan, C. J. *Nature News & Views* **365**, 602–603 (1993).
5. Wambsganss, J. *Nature News & Views* **370**, 248–249 (1994).

ASTROPHYSICS

A Galactic speed record

Galen Gisler

A SOURCE with apparent transverse motion faster than the speed of light has now been detected within our Galaxy, according to I. F. Mirabel and L. F. Rodríguez. Their report appears on page 46 of this issue¹.

At the beginning of the 1970s, apparent faster-than-light motions were reported within the quasars 3C 279 and 3C 273 (refs 2, 3) using Very-Long-Baseline Interferometry (VLBI). Because real faster-than-light motion is impossible according to physics as we know it, these reports were used as 'evidence' for the non-cosmological nature of the redshifts of these objects⁴. The redshift debate then raging made it rational, in some circles, to invoke new physics in the form of a new cause for redshifts, to avoid having to invoke new physics in the form of an explanation for apparent faster-than-light motions.

Illusion

Of course, apparent faster-than-light motion is merely a relativistic illusion that contains no new physics, and this explanation had been made well in advance of these discoveries. As Martin Rees wrote in 1966, "... an object moving relativistically in suitable directions may appear to a distant observer to have a transverse velocity much greater than c " (ref. 5). The usual application of this idea is to a jet beamed towards the observer. The nearer part of the jet is seen at successively later proper times, compared to the core, at successive epochs of observation; the jet's transverse motion therefore appears to be much more rapid than if the near part and the core were seen at the same proper

instant at each epoch of observation.

Since the early 1970s, large numbers of extragalactic sources with superluminal motion, as this phenomenon is now called, have been discovered⁶, and such sources are the strongest evidence we have for the existence of relativistic motions in celestial objects. The multiplier between the true speed and the apparent speed in the standard picture can be as great as the Lorentz factor γ , and the observations imply bulk γ up to 10 in some cases, although the numbers depend on the necessarily uncertain true distances.

The highest intrinsic (which is to say, non-cosmological) velocity directly measured in any object, however, remains the famous 0.26c in the Galactic source SS433 (ref. 7). At γ of about 1.04, this is hardly relativistic. There is a big jump in energetics from this object to quasars, some of which have apparent VLBI superluminal motions. The great distances of these latter sources lead to many uncertainties in the velocity measurements and their interpretations, even notwithstanding the redshift controversy.

A superluminal source in our own Galaxy would be the ideal object for proving the existence of relativistic motion in energetic sources; this philosopher's stone would help bridge the gap between the very energetic jet sources we see at great distances and milder objects such as SS433.

Mirabel and Rodríguez have now handed us such an object. They observed the hard X-ray transient GRS1915+105 with the Very Large Array (VLA) during the early months of 1994. In the course of these observations, a pair of bright con-

densations formed and separated from the compact core. These two condensations moved away at different apparent speeds, with the slower one being fainter. This asymmetry immediately suggests a relativistic interpretation along the classical lines. Using this interpretation, and assuming that the event was symmetrical in the rest frame of the compact core, one can immediately derive an upper limit to the distance of the source (13.7 kpc, placing it firmly within the Galaxy) and a lower limit to the speed of each component (0.323c).

Unfortunately, this source is located very close to the Galactic plane, beyond 20 magnitudes of optical absorption, so that there is little hope of directly obtaining its distance by traditional means. The kinematic H I distance determined by Mirabel and Rodríguez is 12.5 kpc, yielding a component speed of 0.92c directed at an angle of 70° to the line of sight.

Significance

The significance of this discovery must not be underestimated. If these observations hold up, this source will establish the existence of intrinsic relativistic motion in celestial objects in a way that is independent of the controversy over redshifts. This object, being much more energetic than SS433, offers a local analogue to the very much more energetic sources in galactic nuclei, and may help us understand their physics better than we do now. Finally, it offers a new form of astrometry: sources such as this one (and there may well be others) will, for the first time, make possible surveyor-precision distances at Galactic scale. Because the proper-motion asymmetry determines the true transverse speed, measurement of a Doppler-shifted spectral line could determine the angle of the velocity vector to the line of sight and the true distance to the object.

Further VLA observations of this source are called for, to see if there are further outbursts that confirm the relativistic beaming model, and VLBI observations to study the early evolution of an outburst would be helpful as well. Above all, searches should be undertaken for other similar sources that might perhaps be more accessible at other wavelengths. □

Galen Gisler is at Los Alamos National Laboratory, Los Alamos, New Mexico 87544, USA.

1. Mirabel, I. F. & Rodríguez, L. F. *Nature* **371**, 46–48 (1994).
2. Whitney, A. R. *et al.* *Science* **173**, 225–230 (1971).
3. Cohen, M. H. *et al.* *Astrophys. J.* **170**, 207–217 (1971).
4. Stubbs, P. *New Scientist*, **50**, 254–255 (1971).
5. Rees, M. J. *Nature* **211**, 468–470 (1966).
6. Zensus, J. A. & Pearson, T. J. (eds) *Superluminal Radio Sources* (Cambridge Univ. Press, 1987).
7. Margon, B. A. *Rev. Astr. Astrophys.* **22**, 507–536 (1984).

Docking mission accomplished

David Stuart

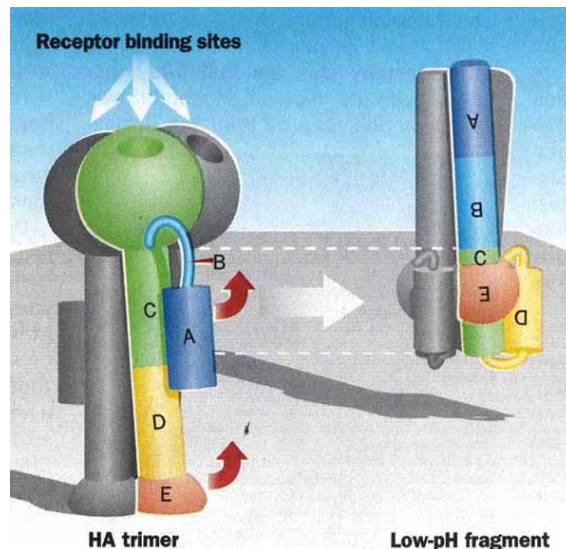
FOR every biological problem that has succumbed to structural analysis, there are many more that have kept their secrets. Often we tactfully gloss over these and turn to easier, if less significant, questions. The alternative, head-on (or heroic) approach is exemplified by a paper that appears on page 37 of this issue¹, which provides dramatic insight into the problem of how proteins cause the controlled fusion of biological membranes — by unprecedented contortions.

Membrane fusion between cells and between intracellular organelles is essential in eukaryotic organisms; to an enveloped virus, fusion of viral and host membranes is vital, opening up a route for its genetic material to enter the cytoplasm of the host cell and use the facilities inside to ensure its own survival. For many years influenza virus has provided the paradigm for this process. The viral protein haemagglutinin (HA) is a central player — each subunit of this glycosylated trimeric molecule starts life as a single polypeptide chain which is cleaved into two, producing the mature molecule which protrudes from the viral membrane ready to bind to a receptor on the target cell. The HA molecule remains attached to the cellular receptor during endocytosis of the virion and then, at the lowered pH required for fusion, undergoes extensive conformational changes, initiating a process that leads to the fusion of the membrane of the virus with that of the target cell.

Thirteen years ago the atomic structure of the soluble extra-virion portion of the mature influenza virus haemagglutinin molecule was published², marking the intrusion of crystallography into the study of animal viruses. This structure, essentially a globular head (comprising most of chain HA₁) mounted on a largely helical stalk (mainly chain HA₂), provided immediate and satisfying insights into the physical basis of antigenic variation³. Subsequently, a sustained attack by the group of Don Wiley has yielded an understanding of how the HA head recognizes a terminal sialic acid residue on the glycoprotein receptor for the virus⁴ and is now guiding the design of anti-influenza virus compounds⁵. But the structure offered no easy answers to the question of how membrane fusion is achieved. A hydrophobic peptide at the amino terminus of HA₂ had been identified as the portion of the polypeptide that

inserts into the target membrane, yet this was 100 Å from the tip of the molecule (and therefore even further from the target membrane), tucked into the stalk of the molecule. In spite of much further work, including elegant mutagenesis studies that revealed that separation of the heads of the trimer is required^{6,7}, no consensus mechanism for fusion emerged.

A way forward was clear: induce the



Schematic representation of the conformations of the HA molecule² and low-pH conformation of a fragment visualized by Bullough *et al.*¹. The HA trimer is shown protruding from a plane representing the surface of the viral membrane. The main structural units of the low-pH structure (from about residues 40 to 155 of the HA₂ chain) are labelled A to E from the N-terminal end and coloured similarly in both structures. The fusion peptide, N-terminal to unit A, would protrude well above the heads bearing the receptor-binding sites, ideally placed to engage the membrane of the target cell.

low-pH conformational changes of HA *in vitro* and determine the structure. Initially this route was blocked by aggregation of the low-pH form, presumably as a result of the exposure of the hydrophobic fusion peptide, frustrating crystallographic analysis. The low-pH form is very susceptible to proteolysis, however, allowing parts of the molecule, including the hydrophobic peptide and most of the HA₁ chain, to be clipped off, leaving the soluble fragment that Bullough *et al.* have now visualized. At first glance the structure (see figure) seems closely related to that of the fragment of the molecule from which it is derived — a trimeric coiled-coil backbone of three helices with a small globular domain at the base and a buttressing α -helix. But look again — the loop preceding the buttressing helix of the high-pH form has become an extension of the coiled coil, as has the buttressing helix,

casting the amino terminus of HA₂ (to which the fusion peptide would be attached) some 100 Å outwards towards the target membrane. This massive rearrangement was not totally unexpected: the residues forming the resultant coiled coil had been implicated in such a structure way back before the structure of HA was known⁸, and experimental evidence for such a rearrangement has recently been produced⁹ (although the structural predictions of this recent work were only partially correct). But the low-pH structural rearrangements don't stop there: at the carboxy-terminal end of HA₂ there is another structural inversion — a piece of the central helix breaks off and folds back to replace (roughly) the helix that has flipped out, and in so doing it takes with it the globular C-terminal subdomain, turning it upside down and apparently disordering the final helix.

These dramatic conformational changes suggest that the energy required for fusion is stored up during the production of the HA molecule. Remember that each of the three subunits starts out as a single polypeptide chain, folded into a stable conformation. On cleavage to form the mature HA found on infectious virus particles, the free ends produced snap 20 Å apart to give a metastable structure. Cleavage has provided the possibility of exploring new structures and the molecule is like a set trap: a reduction in the endosomal pH then springs the trap, setting in motion a series of events that we can perhaps now begin to understand. The fusion peptide is cast out to engage the target membrane (the globular heads separating to make room for this), while tension is generated at the base of the molecule so that it is ready to flip over, dragging the membranes into proximity.

How general are these insights into fusion likely to be? On the one hand, the apparent capacity of evolutionary processes for patiently discovering diverse sol-

1. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. *Nature* **371**, 37–43 (1994).
2. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366–373 (1981).
3. Wiley, D. C., Wilson, I. A. & Skehel, J. J. *Nature* **289**, 373–378 (1981).
4. Weis, W. *et al.* *Nature* **333**, 426–431 (1988).
5. Watowich, S. J., Skehel, J. J. & Wiley, D. C. *Structure* **2**, 719–731 (1994).
6. Godley, L. *et al.* *Cell* **68**, 635–645 (1992).
7. Kemble, G. W., Bodian, D. L., Rose, J., Wilson, I. A. & White, J. M. *J. Virol.* **66**, 4940–4950 (1992).
8. Ward, C. W. & Doppeide, T. A. *Aust. J. Biol. Sci.* **33**, 441–447 (1980).
9. Carr, C. M. & Kim, P. S. *Cell* **73**, 823–832 (1993).
10. Flore, O., Fricks, C. E., Filman, D. J. & Hogle, J. M. *Sem. Virol.* **1**, 429–438 (1990).
11. Goldsmith, E. J. & Mottonen, J. *Structure* **2**, 241–244 (1994).

utions to biological problems and the relative simplicity of the fusion mechanism in influenza virus suggest that other systems may employ different tactics. On the other, the strategy of using proteolytic cleavage to set a molecular trap seems to have found more general application. Thus many viruses produce precursor polypeptides which are cleaved to form the mature polypeptides, often with substantial separation of the freed ends. An

example of this is found in the Picornaviridae which, in spite of possessing no lipid and hence having no need of membrane fusion to get their genetic material into the cell, appear to drive infection in an analogous way. In this case, cleavage of the polypeptide chains during maturation of the capsid seems to prime the virus for pH- or receptor-mediated extrusion of a hydrophobic structure into the host cell membrane, initiating release of the viral

RNA¹⁰. Indeed, like many really useful biological engines, this one pops up in disparate contexts, for instance in what is perhaps the most dramatic irreversible structural rearrangement yet seen among eukaryotic protein molecules, the serpin family of protease inhibitors¹¹. □

David Stuart is in the Laboratory of Molecular Biophysics, Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK.

OBITUARY

Dorothy Hodgkin (1910–94)

DOROTHY Hodgkin died on 29 July, thirty years after her receipt of the Nobel Prize for Chemistry "for her determination by X-ray techniques of the structures of biologically important molecules". Each of her structures extended X-ray crystallography to molecules of greater complexity than any previously analysed. She established it as one of the fastest methods of finding the chemical constitution of natural products; nuclear magnetic resonance has since become the other.

In 1943, Dorothy solved her first structure, that of cholesterol iodide. She assumed that the scattering contributions of the heavy iodine atoms outweighed the combined ones of all the light atoms and hence determined the phases; but because the two iodines in the unit cell are related by a centre of symmetry, and the molecule contains chiral centres, this gave the right structure superimposed on its mirror image. Undeterred, she devised an ingenious way of distinguishing the correct from the spurious peaks in the electron density map, and the structure fell out. She was disappointed when it added merely a few stereochemical details to the formula of cholesterol already elucidated by the chemists.

Penicillin, first isolated in 1942 by Ernst Chain a stone's throw away from Dorothy's Oxford laboratory, presented a greater challenge because its chemical constitution was then still unknown. Initially the crystals proved useless, and she had to wait until February 1944 when the first good ones arrived from the United States. Fourier projections of the sodium salt phased by the heavy atom method, and of the rubidium and potassium salts phased by isomorphous replacement, gave only the vaguest outlines of the molecule. Again, Dorothy refused to be discouraged. By a remarkable piece of chemical intuition, she inferred the right answer and proved it by an elegant three-dimensional Fourier synthesis which revealed the complete stereochemistry, including the very unusual β -lactam ring proposed a few weeks earlier by E. P. Abraham.

The next challenge was vitamin B₁₂, which contains more than four times as many atoms as penicillin. But in 1948, when Dorothy took the first X-ray pictures, even its molecular weight was unknown. She records that "the discov-

rendered the determination of heavy atom positions needed for isomorphous replacement extremely difficult. Dorothy was delighted when, in 1988, a genetically modified human insulin promised an improved treatment for diabetes. This could not have been designed without the structure on which she had laboured for most of her active life.

Dorothy Hodgkin radiated love: for science, her family, her friends, her students and her crystals. That love was combined with a brilliant mind and an iron will to succeed. Her uncanny knack of solving structures owed much to her wide knowledge of stereochemistry (she used to write the chapters on crystal structures for the Annual Reports of the Chemical Society and kept all their stereochemistry in her head). She and her students had to solve their structures laboriously by taking series of Weissenberg photographs with weak X-rays from sealed tubes, by estimating the intensities of hundreds or even thousands of spots visually, and by calculating structure factors and initially even Fouriers on manually operated calculators. Any of these structures could now be solved in a few days by feeding X-ray data collected on computer-controlled diffractometers into other computers programmed to find atomic positions automatically by direct methods. Even Dorothy's favourite molecule, insulin, with a molecular weight of 5,778, is now beginning to yield to direct methods, and may also soon be solved automatically. She would have rejoiced at that phenomenal progress of X-ray analysis.

Science apart, Dorothy's life was devoted to many good causes. But I thought that in this memoir I should concentrate on her research rather than her work for peace, for the Third World, or the students of Somerville College and of Bristol University of which she was Chancellor. In *Nature* she would have wanted to be remembered above all as the superb scientist she was.

Max Perutz

Max Perutz is in the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Dorothy Hodgkin — a brilliant mind and an iron will to succeed.

ery of the cobalt atom... first stimulated us to attempt a detailed investigation". Dorothy solved its structure and discovered its remarkable chemical constitution by her imaginative, bold interpretation of three-dimensional Patterson functions and of blurred electron-density maps phased on the cobalt atom alone, ingeniously picking out the significant peaks from a jungle of spurious ones. To succeed, she had to shed the obvious assumption that the groups surrounding the cobalt were pyrrole rings as in porphyrins and discovered the quite novel ring structure which she named corrin.

In 1935, at the very start of her career, Dorothy had prepared her first crystals of insulin and discovered their rich X-ray diffraction pattern which made her determined to solve its structure. She persevered with her attempts in parallel with all her other work. To her disappointment, however, insulin did not yield until 1969, mainly because it crystallized in the rhombohedral space group R3 which lacks centrosymmetric projections and

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

p53, guardian of Rb

Eileen White

REGULATION of cell proliferation, an activity subverted by the loss of tumour-suppressor gene function, is a key factor in the development of cancer. Over the past two years it has gradually emerged that the control of cell proliferation is intimately connected to the control of programmed cell death or apoptosis. Two reports — one in *Genes and Development*¹, one on page 72 of this issue² — now show that, in the developing ocular lens, loss of function of the retinoblastoma (Rb) tumour-suppressor gene releases a growth restriction, permitting cellular DNA synthesis during terminal differentiation of otherwise post-mitotic lens fibre cells. Growth deregulation is counteracted by apoptotic cell death orchestrated by yet another product of a tumour-suppressor gene, p53. So the loss of one tumour-suppressor gene is compensated by the activity of another, serving as a safeguard mechanism to control the emergence of neoplastic cell growth.

Indications that Rb protein may inhibit apoptosis came from the observation that the DNA-virus transforming proteins adenovirus E1A and human papilloma virus (HPV) E7, which inhibit Rb function as a mechanism to deregulate cell growth, also induce apoptosis³⁻⁵. Furthermore, transgenic mice with homozygous inactivation

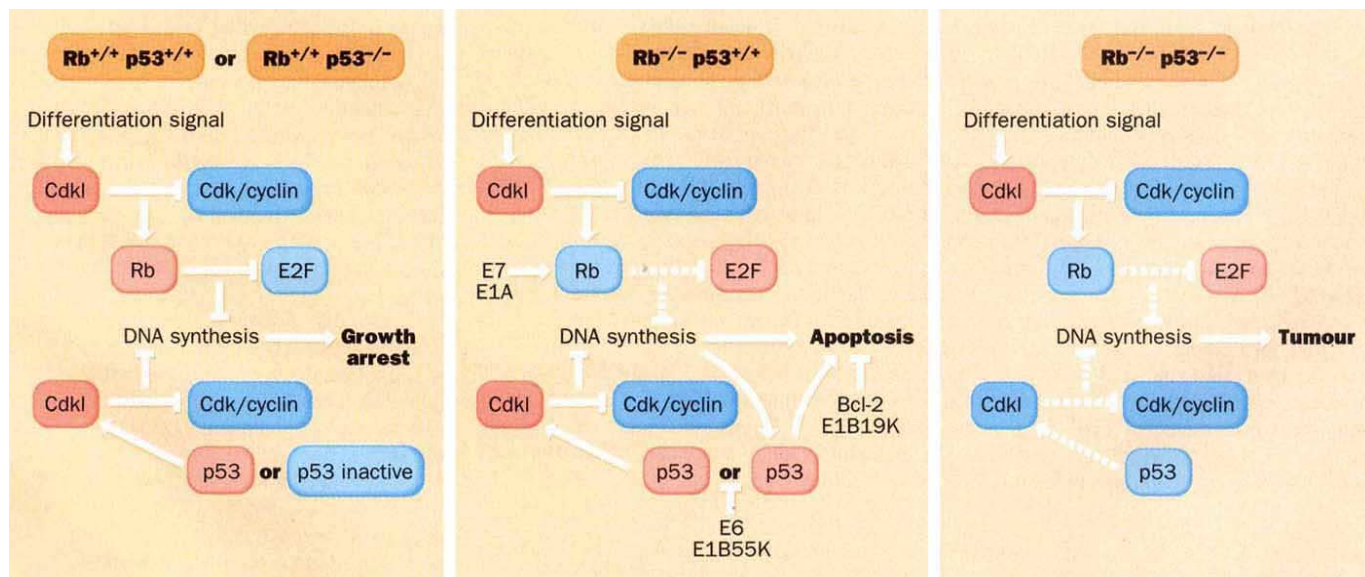
of Rb (*Rb*^{-/-}) display widespread cell death in the nervous system and die mid-gestationally⁶⁻⁸.

Pan and Griep¹ applied the first approach using targeted HPV E7 gene expression to inactivate Rb (and its relatives *p107* and *p130*) in lens fibre cells, whereas Morgenbesser and colleagues² examined differentiation of lens-fibre cells in *Rb*^{-/-} mouse embryos. In both cases, interference in Rb function had profound effects on lens development. Lens cells that would normally withdraw from the cell cycle, elongate and differentiate into fibre cells, instead continue to synthesize DNA and die by apoptosis. Disorganization of fibre cells occurs in embryos, and adult mice with the E7 ablation of Rb function display profound microphthalmia (small eyes). Thus, Rb implements the cell-cycle arrest required for normal differentiation, and continued progression of the cell-cycle is incompatible with cell survival.

The Rb protein is an inhibitor of cell-cycle progression at G1/S by virtue of its ability to inhibit the activity of members of the E2F family of transcription factors⁹. The presence of E2F sites in the promoters of genes encoding thymidine kinase, Myc, Myb, dihydrofolate reductase and DNA polymerase- α , the prod-

ucts of which are implicated in induction of S phase, suggests that E2F family members may be responsible for traversing the G1/S restriction point of the cell cycle. Indeed, E2F1 overexpression is sufficient to promote DNA synthesis in growth-arrested cells¹⁰, but also drives cells into apoptosis¹¹. Inappropriate or untimely DNA synthesis produced by the loss of Rb and the activation of E2F may initiate the apoptotic response. The question remained as to the cellular mechanism responsible for implementing the cell-death response to the loss of Rb.

There is increasing evidence that another tumour-suppressor protein, p53, may induce apoptosis. Overexpression of p53 can induce either growth arrest or apoptosis, depending on the cell type¹². Conversely, loss of p53 function in mice that have homologous inactivation of *p53* (*p53*^{-/-}) can produce resistance to apoptotic cell death^{13,14}. The idea that apoptosis resulting from Rb loss in the lens may be mediated by p53 derived from two observations: first, that human tumours have often lost both *p53* and *Rb*; second, that DNA-virus transforming proteins target both *Rb* and *p53* for inactivation. In the case of adenovirus E1A, which inactivates Rb and induces apoptosis, induction of apoptosis depends on functional p53 and is blocked by gene products that interfere directly and indirectly with p53 function^{15,16}. From the new experiments involving both gene knock-out² and targeted HPV gene expression¹, it seems



Balance and counterbalance: control of cell proliferation and apoptosis by Rb and p53. In normal cells (*Rb*^{+/+}, *p53*^{+/+}), cell proliferation can be inhibited through the activation of Cdkls and functional sequestration of E2F by Rb. By transcriptionally activating a Cdkl, p53 can also inhibit cell-cycle progression, or it may be inactive which does not further alter the growth-arrested state. p53 function is not required as long as Rb function is intact (*Rb*^{+/+}, *p53*^{-/-}). Rb loss with active p53 (*Rb*^{-/-}, *p53*^{+/+}) permits E2F-dependent stimulation of DNA synthesis. E7 and E1A interact with Rb at the protein level to activate E2F. Growth arrest imposed by p53 may conflict with induction of DNA synthesis by active E2F, and the opposing growth signals may be mani-

festated by the induction of apoptosis. Alternatively, p53 may become activated in response to Rb loss and induce apoptosis by a cell-cycle-independent mechanism. The Bcl-2 and E1B 19K transforming proteins block p53-dependent apoptosis indirectly, whereas HPV E6 and E1B 55K proteins inhibit p53 directly. Loss of both Rb and p53 function (*Rb*^{-/-}, *p53*^{-/-}) permits cell proliferation without apoptosis, ultimately resulting in unrestricted growth and tumorigenesis. Solid lines, active pathways; broken lines, inactive pathways. Arrows and bars indicate stimulatory and inhibitory events, respectively. Red, active components of the pathways; blue, inactive components.

that apoptosis resulting from *Rb* inactivation in the lens also depends on functional p53.

Embryos that are genotypically *Rb*^{-/-}, *p53*^{-/-}, or those doubly transgenic for E7 and E6 expression in the lens (to abrogate *Rb* and p53 function, respectively), display dramatically reduced apoptosis in lens fibre cells^{1,2}. Fibre cells continue to proliferate because they are released from the *Rb* growth restriction, suggesting that p53 is required as a tumour suppressor to respond to aberrant loss of *Rb* by implementing cell death.

Because of the embryonic lethal phenotype, the issue of whether continued cell proliferation in the absence of apoptosis leads to increased tumorigenicity cannot be addressed in *Rb*^{-/-} mice. Adult mice with lens-specific E7 and E6 expression, however, develop lens tumours, whereas mice transgenic for E7 and E6 alone, do not¹. The cooperativity between the loss of *Rb* and *p53* in oncogenesis can be explained by the requirement that cell proliferation and apoptosis accompanying *Rb* loss must be coupled to the loss of the apoptotic activity of p53. Cell proliferation can be sustained only in the absence of the p53 safeguard mechanism to eliminate cells engaged in untimely DNA synthesis or cell-cycle progression during terminal differentiation. With p53 function intact, the loss of *Rb* is held in check by the apoptotic response mechanism controlled by p53.

On the one hand, it is understandable that deregulated cell proliferation upon *Rb* loss during terminal differentiation would have profound effects in development⁶⁻⁸, as *Rb* is apparently essential for exit from the cell cycle. On the other, the loss of the p53 safeguard alone is not immediately detrimental to the differentiation process¹⁷ until a second event (such as an *Rb* mutation) occurs, in which case aberrant cell proliferation can be tolerated and tumour formation initiated.

As the mechanism of regulation of apoptosis becomes clearer, it is interesting to note that *Rb* and p53 both interact functionally at the level of the cell cycle (see figure). Control of E2F by *Rb* can regulate passage from G0/G1 and into S phase¹⁰. Transcriptional activation of the

WAF1 (also known as p21, Cip1 and Sdi1) cell-cycle-dependent kinase inhibitor (CdkI) by p53 can inhibit several cell-cycle-dependent kinases (Cdk) and cell-cycle progression¹⁸⁻²¹. With *Rb* function intact, cell proliferation can be arrested in differentiating cells by functional sequestration of E2F by *Rb*. As p53 can promote growth arrest as well, its presence may be redundant and therefore not apparent. p53 may also be inactive as long as *Rb* function is intact. In the absence of *Rb*, created by genetic loss of *Rb* gene function or through the expression of E7 or E1A, which interact with *Rb* at the protein level, E2F activation stimulates cellular DNA synthesis.

p53, however, can inhibit cell-cycle progression through CdkI inhibition of other Cdk/cyclin complexes, and this failure to stop growth properly as a result of conflicting growth signals may produce apoptosis. Alternatively, p53 may become activated in response to *Rb* loss, perhaps through the generation of DNA damage, and implement apoptosis independently of transcriptional transactivation of cell-cycle regulatory genes²². The HPV E6 and adenovirus E1B 55K proteins can block apoptosis in the situation of *Rb* loss and p53 activation through direct interference with p53 (refs 4, 5), whereas Bcl-2 and the adenovirus E1B 19K protein override apoptosis indirectly^{16,15}. Without both *Rb* and p53, E2F activation stimulates cell proliferation, which is unobstructed in the absence of p53, permitting tumour formation.

Although it is attractive to link cell-cycle control and apoptosis because a checkpoint for abnormal cell growth is then automatically in place, it is still possible that p53 may implement the apoptotic programme by a mechanism that has nothing to do with the cell cycle. Nonetheless, the involvement of tumour suppressors in the regulation of apoptosis in normal differentiation is now firmly established and should provide fertile ground for future work. □

Eileen White is at the Center for Advanced Biotechnology and Medicine, and the Department of Biological Sciences, Rutgers University, 679 Hoes Lane, Piscataway, New Jersey 08854, USA.

Light and shade

For many white-skinned people, a suntan is the height of chic. Sadly, it has a fearsome sting at the end of a very long tail. The same solar ultraviolet light that provokes the tan also slowly attacks and degrades the collagen fibres of the dermis beneath it. After many years of exposure, the dermis loses its taut elasticity, and folds into wrinkles. Some people are so afraid of facial wrinkles that they try to avoid sunlight altogether. They wear sunblock creams all the time, summer and winter, to screen out the ultraviolet and stave off the encroaching creases. Daedalus is now resolving the whole dilemma.

He points out that many substances absorb light and re-emit it at twice the frequency. Some are natural organics, such as urea and β -carotene, but certain 5-nitropyridine derivatives show the effect more strongly. For good results the molecules must be correctly ordered: either in a crystal lattice or within a Langmuir-Blodgett film. So DREADCO cosmetics are formulating a skin cream around lipid-soluble 5-nitropyridine derivatives. It will carry the second-harmonic generator molecules into the epidermis, where they will be absorbed and aligned by the ordered lipid membranes of the melanocyte cells.

The user will then be able to tan in ordinary daylight. The second-harmonic generator will absorb this light, and fluoresce at twice the frequency — neatly in the ultraviolet. The weak fluorescence, generated within the melanocytes themselves, will trigger them as effectively as the fiercest solar ultraviolet beating down from outside. They will synthesize melanin copiously to defend the skin against this seemingly ferocious radiative assault. In fact the stimulation will be entirely local, and no other skin structures will be irradiated at all.

DREADCO's 'Fluorotan' cream will provoke a tan in the duldest or most overcast weather, or even in a British November fog. It will act as a sort of pre-emptive inoculation against solar ultraviolet light. When real sunlight comes along, the user will be pre-tanned, indeed over tanned, for the occasion. His epidermis, already fully charged with skin-defending melanin, will ward off both the threat of skin cancer and the wrinkles of old age.

Occasional reapplications of Fluorotan will keep this smart, safe, protective tan permanently in place. It will even work under sunblock cream (which is transparent in the visible). Elegant in youth, and safe from the corrugations of age, the user will enjoy the best of both worlds. David Jones

- Pan, H. & Griep, A. E. *Genes Dev.* **8**, 1285–1299 (1994).
- Morgenbesser, S. D., Williams, B. O., Jacks, T. & DePinto, R. A. *Nature* **371**, 72–74 (1994).
- White, E. *et al. Molec. cell. Biol.* **12**, 2570–2580 (1992).
- Rao, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 7742–7746 (1992).
- White, A., Livanos, E. M. & Tlsty, T. D. *Genes Dev.* **8**, 666–677 (1994).
- Lee, E. Y., H. P. *et al. Nature* **359**, 288–294 (1992).
- Jacks, T. *et al. Nature* **359**, 295–300 (1992).
- Clarke, A. R. *et al. Nature* **359**, 328–330 (1992).
- Nevins, J. R. *Science* **258**, 424–428 (1992).
- Johnson, D. J., Schwarz, J. K., Cress, W. D. & Nevins, J. R. *Nature* **365**, 349–352 (1993).
- Wu, X. & Levine, A. J. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3602–3606 (1994).

- Yonish-Rouach, E. *et al. Nature* **352**, 345–347 (1991).
- Clarke, A. R. *et al. Nature* **362**, 849–852 (1993).
- Lowe, S. W., Schmitt, E. M., Smith, S. W., Osborne, B. A. & Jacks, T. *Nature* **362**, 847–849 (1993).
- Debbas, M. & White, E. *Genes Dev.* **7**, 546–554 (1993).
- Chiou, S.-K., Rao, L. & White, E. *Molec. cell. Biol.* **14**, 2556–2563 (1994).
- Donehower, L. A. *et al. Nature* **356**, 215–221 (1992).
- El-Deiry, W. S. *et al. Cell* **75**, 817–825 (1993).
- Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. *Cell* **75**, 805–816 (1993).
- Noda, A., Ning, Y., Venable, S. F., Pereira-Smith, O. M. & Smith, J. R. *Exp. Cell Res.* **211**, 90–98 (1994).
- Xiong, Y. *et al. Nature* **366**, 701–704 (1993).
- Caelles, C., Helmborg, A. & Karin, M. *Nature* **370**, 220–223 (1994).

The 'St Jude' mind virus

SIR — We have previously suggested the existence of human information parasites, or, in their simple form, "viruses of the mind"¹⁻³. The St Jude chain letter (reproduced, right) provides an example. It is simple, direct and, by its apparent longevity, very fit. Anecdotal data⁴⁻¹² confirm its virulence and suggest that it is one of a class of postal parasites.

A virus is a piece of code that promotes its own replication. Viruses are parasitic in that the energy and other costs of duplication are borne by the hosts and not by the virus. Conventional viruses propagate DNA or RNA information, through cellular machinery that is already set up to obey their language. Computer viruses succeed because computers are set up slavishly to obey the programming languages in which their 'duplicate me' instructions are written. 'Viruses of the mind' or 'memes'¹⁻³ would also succeed in getting themselves duplicated if human minds were set up to obey them. A piece of paper bearing the words 'Make 10 copies of me and send them to 10 people' would spread its message like a brushfire, if only brains obeyed instructions as slavishly as computers. There seems no obvious reason, however, why people should be so mindlessly obliging.

The phenomenon of the chain letter is a sobering case. (The name is not apt because a chain is normally a linear array of links; the 'chain' letter, on the other hand, forms an exponentially branching tree.) A particularly stark example, the St Jude letter, was received recently by Alison Clarkson, who showed it to her husband (O.R.G.). O.R.G. promptly sent it to R.D. The analogy to biological viruses is clear. St Jude infects potential hosts through the vector of the post. By inducing guilt, fear, greed and piety, it causes susceptible hosts to multiply it 20-fold and transmit the 20 copies to new potential hosts through the postal vector. Each new host then potentially initiates the production of another 20 copies. Whether or not any particular infection is successful, St Jude's hosts can suffer mental distress, as real, in its own way, as the physical distress caused by the common cold virus.

There is anecdotal evidence for the longevity and distribution of the St Jude strain of postal virus. By its own account, which must be viewed with some scepticism, the St Jude strain was in circulation as early as 1903. Paul M. Griffo, national spokesman for the US Postal Inspection Service (USPIS), said in a recent interview: "Ah, the St Jude letter. Nine is a low estimate of how many times that thing has been around the world. It's as old as dirt."⁴ Griffo confirms (personal communication) that it goes back farther than the institutional memory of the US postal

service, and has periodic outbreaks. He reports that the St Jude virus has had various incarnations — for example the "A.R.P." officer in the text of the letter (right) is, in some other versions, referred to as an RAF officer. The present outbreak (which we refer to as St Jude 1) has included members of the news media among its sufferers, with the result that the recent appearance of St Jude 1 or a near variant can be documented or inferred in locations as far apart as Seattle⁴, England⁵ and Dallas⁶.

Many potential hosts must be immune. If all hosts were susceptible to each infection, each successive generation of St Jude 1 would grow by a power of 20. By the end of eight generations, there would be a total of 20^8 or 2.56×10^{10} copies in the post. At this rate, every one in the world, on average, would receive approximately 4.5 copies. The fact that this has not occurred is negative evidence of immunity. Previous infection with any kind of chain letter is likely to raise resistance. Furthermore, St Jude 1 is perhaps counterproductive in specifying 20 copies. Many would-be obeyers might be put off by the labour of making and sending as many as 20 copies. Twofold duplication might, paradoxically, show more effective penetration.

We turned out to be immune, although we both admit to experiencing waves of mild, irrational anxiety on deciding not to comply, and we could be said to seek a modicum of good luck by sharing it, on a purely scientific basis, through the medium of this journal. Indeed, this report constitutes a form of meme therapy, for we have attached our own information package to a mind virus of proven virulence.

Other strains of postal virus use slightly different manipulative techniques to engineer their propagation. Some promise money to their hosts⁴. Examples of this have recently appeared on the Internet⁷. Other chain letters currently circulating involve the exchange of women's

With Love All Things are Possible.

This paper has been sent to you for Luck. The original is in New England. It has been sent around the world. The Luck has been sent to you. You will receive good luck within 4 days of receiving this letter pending in turn you send it on. This is no joke. You will receive good luck in the mail. Send no money. Send copies to people you think need good luck. Do not send money cause faith has no price. Do not keep this letter. It must leave your hands within 96 hrs. An A.R.P. officer Joe Elliot received \$40,000,000. George Welch lost his wife 5 days after this letter. He failed to circulate the letter. However before her death he received \$7,775,000. Please send copies and see what happens after 4 days. The chain comes from Venezuela and was written by Saul Anthony Degnas, a missionary from S.America. Since that copy must tour the World. You must make 20 copies and send them to friends and associates after a few days you will get a surprise. This is love even if you are not superstitious. Do Note the following: Contonare Dias received this letter in 1903. He asked his Sec'y. to make copies and send them out. A few days later he won a lottery of 20 million dollars. Carl Dobbitt, an office employee received the letter + forgot it had to leave his hands within 96 hrs. He lost his job. After finding the letter again he made copies and mailed 20 copies. A few days later he got a better job. Dolan Fairchild received the letter and not believing he threw it away. 9 days later he died. In 1987 the letter was received by a young woman in Calif. It was faded and hardly readable. She promised her self she would retype the letter and send it on but, she put it aside to do later. She was plagued with various problems, including expensive car problems. This letter did not leave her hands in 96 hrs. She finally typed the letter as promised and got a new car. Remember send no money. Do not ignor this — it works.

St. Jude

St Jude's letter (above). Small gaps left by repeated photocopying of the hand-written original have been interpolated; otherwise, the text is a verbatim transcription.

underwear^{5,8}, postcards of "naked Asian Girls"⁹, and protests on the disappearance of a young woman¹⁰. It has been reported that one particularly virulent chain, which originally requested get-well wishes and/or business cards for a young cancer patient in the UK named Craig Shergold, has produced more than 70 million responses¹¹. Despite published pleas to stop, "mountains of unwanted business cards" keep arriving daily¹². The Craig Shergold letter has been documented in Hong Kong, where the governor, Chris Patten, is reported to have participated⁹.

St Jude 1 provides confirmation for the existence of human mind viruses. It is, of course, still a leap to a more general theory of more complicated mental parasites and symbionts. It can be argued that crude viruses like St Jude are simply doing on a small, crude, but robust level what culture systems are doing more pervasively and with greater complexity. Whether

1. Dawkins, R. in *Dennett and his Critics* (ed. Dalhborn, B.) 13-27 (Blackwell, Oxford, 1993).
2. Goodenough, O. in *Proceedings of the London School of Economics Conference on Evolution and the Human Sciences*, 12 (LSE, London, 1993).
3. Dawkins, R. *The Selfish Gene* (Oxford Univ. Press, Oxford, 1976).
4. Foster, D. *Los Angeles Times Metro*, Part B, p. 6. (13 February 1994).
5. Bresler, F. *The Daily Telegraph*, p. 6 (26 March 1994).
6. Powell, L. *The Dallas Morning News*, p. 18A (29 March 1994).
7. Burgess, J. *The Washington Post*, Financial, p. F19 (28 February 1994).
8. *Press Association News File*, Home News (6 March 1994).
9. *South China Post*, Business, p. 16 (14 February 1994).
10. Adcock, S. *Newsday*, News, p. 4 (10 March 1994).
11. Downey, M. & Shurling, B. *The Atlanta Constitution*, Section E., p. 2 (28 February 1994).
12. Winter, C. *Chicago Tribune*, Metro Northwest, p. 1 (20 April 1994).

or not this leap is justified, the identification of St Jude 1 helps to move the debate about viruses of the mind from the abstract to the concrete and may lead to the identification of other examples.

We have no intention of launching experimental chain letters into the general population for the purpose of testing hypotheses of quantitative epidemiology, for measuring mutation rates, or for assessing the limits of human gullibility.

Oliver R. Goodenough

Vermont Law School,
South Royalton,
Vermont 05068, USA

Richard Dawkins

Department of Zoology,
University of Oxford, South Parks Road,
Oxford OX1 3PS, UK

Pulsed dynamics of fountains

SIR — The fascination with ornamentation fountains probably goes back to the *quattrocento* of Francesco di Giorgio and later Leonardo da Vinci^{1,2}. The unsteady character of a fountain's flow can reinforce its attraction, and several designs have exploited the intermittency of the flow — for example the so-called Hero fountains³ such as the *Fontana a tempo* of

di Giorgio², or spontaneously fluttering water sheets^{4,5}. Here I show that a vertical water jet undergoes self-sustained pulsations in height, associated with the gravity-induced backflow of the jet on itself.

Consider the first fluid element to emerge from a vertical fountain. The maximum height that it can reach (assuming no frictional losses with the surrounding air) is given by $h = u_0^2/2g$, where u_0 denotes the velocity with which the fluid emerges and g is the acceleration due to gravity. Having reached the maximum height, the fluid element will start to fall, accelerating downwards under the influence of gravity. This gives rise to a backflow in the jet, which initially takes the form of a stationary 'cap' — a fluid packet perched on top of the jet and fed with liquid from below. It then falls under its own weight, flattening out the ascending column of the fluid while deforming under the inertial pressure of the jet. The cap eventually breaks up into a dispersed corolla, the jet re-emerges and a new cycle begins. The overall effect is a pulsating motion of the jet (Fig. 1), which is apparent not only to the eye but also to the ear (the break-up of the corolla is accompanied by a characteristic dripping sound). It can be easily shown that the pressure at the orifice of the jet is related to its instantaneous height. The (mean sub-

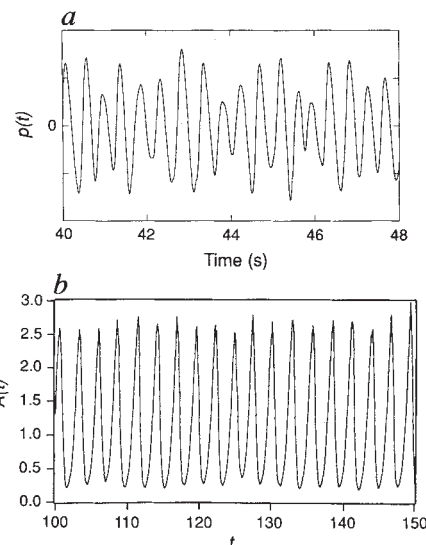


FIG. 2 *a*, Fluctuation of the pressure at the nozzle of the fountain shown in Fig. 1. *b*, Solution for the amplitude $A(t)$ of the nonlinear delayed model equation with $r = 2$, $\tau = 0.6$, $\mu = 1$. The delay has been allowed to fluctuate at each time step: $\tau(t) = \tau + \delta\tau(t)$. The fluctuating part of the delay $\delta\tau(t)$ is distributed according to a uniform probability density function whose width is such that $|\delta\tau|/\tau = 1/2$.

tracted) temporal variation of this pressure displays a dominant period (Fig. 2*a*). For fountains with moderate aspect ratios (height/diameter, $h/d \leq 50$), this oscillatory behaviour develops before the onset of the Rayleigh capillary break-up instability, which would otherwise cause the liquid column to fragment.

The role of gravity-induced backflow in this oscillatory behaviour is readily demonstrated. If a flat, horizontal object is used to obstruct the top of the liquid column, the fluid spreads outwards towards the edges of the obstruction, before falling as droplets away from the main jet (Fig. 1). The backflow is thus removed and the oscillations disappear. Similarly, if the fountain is oriented slightly away from the vertical, backflow is no longer possible and the jet describes a parabola with a fixed maximum elevation.

It has been argued^{6,7} that for unsteady recirculating flows the existence of oscillations derives from the interplay between linear growth and delayed nonlinear saturation^{6,7}. This mechanism can be represented as an evolution equation for the amplitude $A(t)$ of the disturbances in the flow, equivalent to the fluctuating height of the fountain:

$$\frac{d}{dt} A(t) = rA(t) - \mu|A(t - \tau)|^2 A(t)$$

where τ is the transit time through the recirculation loop. This time-lag represents the interaction time of a fluid packet initially topping the fountain during its deformation until its break-up in dispersed droplets. It is of the order of the time for the packet to fall by a distance of

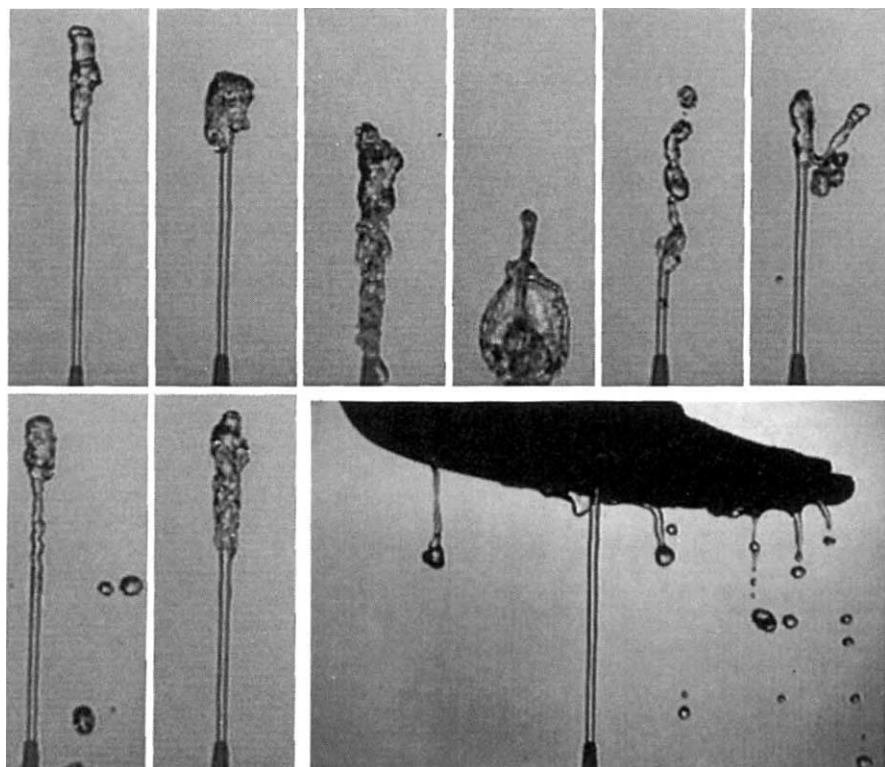


FIG. 1 From left to right and from top to bottom: a period of the pulsation of a vertical fountain. $d = 3$ mm, $u_0 = 1.5$ m s⁻¹. The pictures are spaced by 2/25 seconds and the oscillation frequency is about 2 Hz. Note the corolla break-up of the gravity-induced backflow. The Weber number $We = (\rho u_0^2 d)/\sigma$ is 96, the Froude number $Fr = u_0^2/(gd)$ is 76 and the Bond number $Bo = (\rho g d^2)/\sigma$ is 1.26. Bottom right: Suppression of the oscillations by deflection of the gravitational backflow.

the order of its own size, fixed by the diameter d of the jet: $\tau \sim d/(gd)^{1/2} = (d/g)^{1/2}$, independent of u_0 .

When only one retarded time is involved and when the amplitude $A(t)$ and the delayed feedback $|A(t-\tau)|^2$ are sufficiently out of phase (that is, when $\tau > \pi/4$), the solutions for $A(t)$ are periodic, nonlinear oscillations (Fig. 2b), whose period⁷ is an increasing function of τ . The possible variation of $\tau(t)$ at each instant of time around a mean delay τ may reflect the variability of the size of the packets and of their interaction time in the actual situation. In that case, neither the oscillatory behaviour nor the period is appreciably altered⁷.

Emmanuel Villermaux

Institut de Mécanique
de Grenoble/LEGI—CNRS,
BP 53X, 38041 Grenoble Cedex,
France

1. Di Giorgio Martini, F. *Trattati di Architettura Ingegneria e Arte Militare* (Italy, about 1490).
2. Pedretti, C. *Léonard de Vinci Architecte* (Electa Moniteur, Milan-Paris, 1978).
3. Nollé, A. J. A. *Leçons de Physique Expérimentale III* (Paris, 1743).
4. Casperson, L. W. J. *appl. Phys.* **74**, 4894–4898 (1993).
5. Bousasse, H. *Jets, Tubes et Canaux* (Delagrave, Paris, 1923).
6. Villermaux, E. & Hopfinger, E. J. *J. Fluid Mech.* **263**, 63–92 (1994).
7. Villermaux, E. & Hopfinger, E. J. *Physica D* **72**, 230–243 (1994).

Climate and food supply

SIR — Rosenzweig and Parry^{1,2} have performed a commendable pioneer study of the security of the global food supply in a changing climate. But it is important to note the assumptions and caveats in this study, and the possible effects which it did not consider. These limitations mean that we cannot share the optimistic interpretations put on the findings by Reilly in *News and Views*³. Climate change cannot yet be dismissed as having only minor effects on future global food security.

(1) Rosenzweig and Parry's data suggest that climate change will significantly increase the present unequal distribution of food supplies, thereby exacerbating the problem.

(2) As Rosenzweig and Parry point out, the relatively benign overall result "depends strongly on the full realization in the field of beneficial direct physiological CO₂ effects on crop growth and water use as currently measured in experimental settings". Without these assumed beneficial effects, they find that world cereal production for an equivalent doubling of CO₂ is reduced by 11–20%. The assumed direct effects of CO₂ increase on global wheat yield is about 20 per cent in all three climate change scenarios (Table 9 of ref. 2). Recent studies^{4,5}, including our own⁶, suggest that competitive effects in plant

canopies, as well as other factors, may well result in realized beneficial effects on yield considerably less than in the laboratory.

(3) The ability to make large beneficial adaptations, which Reilly claims could be even greater than assumed by Rosenzweig and Parry, may in fact be exaggerated, and very costly. Rosenzweig and Parry take no account of water-supply limitations for irrigation. Reilly points to the 'green revolution' and the high yields achieved by US agriculture, but these have both relied largely on massive inputs of energy through irrigation, artificial fertilizers and pesticides. One of Rosenzweig and Parry's most striking findings is that even massive adaptations make relatively little difference to the global picture (much smaller than the assumed benefits of increased CO₂).

(4) The climate change factors considered by Rosenzweig and Parry necessarily exclude many potentially important, perhaps even dominating, effects. Many phenomena are not yet reliably represented in global climate models, for example the behaviour of the El Niño Southern Oscillation (ENSO), or changes to tropical cyclone intensity, frequency or location⁷; widespread increases in rainfall intensity, with the corresponding reduction in the average interval between flood events⁸; changes in climatic variability⁹, together with changes in the frequency of occurrence and distribution of pests and diseases; as well as other considerations.

(5) As demonstrated in Fig. 4 of ref. 2, global crop yields are a non linear function of increasing temperature, with substantial gains from a 2°C warming turning to a loss for a 4°C warming. Without substantial greenhouse-gas emission reductions, warming will continue long after the year 2060 (ref. 10), so that a critical threshold (when production will decline in temperate as well as in tropical countries) may well be reached at a later date.

A. Barrie Pittock

Peter Whetton

Yingping Wang

Climate Impact Group,

CSIRO Division of

Atmospheric Research,

Private Bag 1, Mordialloc 3195,

Australia

1. Rosenzweig, C. & Parry, M. L. *Nature* **367**, 133–138 (1994).
2. Rosenzweig, C., Parry, M. L., Fischer, G. & Frohberg, K. *Climate Change and World Food Supply*, Res. Rep. No. 3. (Environmental Change Unit, Univ. Oxford, 1993).
3. Reilly, J. *Nature* **367**, 118–119 (1994).
4. Baker, J. T., Allen, L. H. J. & Boote, K. J. *Agric. For. Met.* **60**, 153–166 (1992).
5. *Austr. J. Bot.* **40**, Nos. 4 & 5 (1992).
6. Wang, Y. P., Handoko, J. R. & Rimington, G. M. *Clim. Res.* **2**, 131–149 (1992).
7. Pittock, A. B. *Weath. Clim.* **12**, 17–31 (1992).
8. Whetton, P. H., Fowler, A. M., Haylock, M. R. & Pittock, A. B. *Clim. Change* **25**, 289–317 (1993).
9. Mearns, L. O., Rosenzweig, C. & Goldberg, R. *Agric. For. Met.* **62**, 159–189 (1992).
10. Manabe, S. & Stouffer, R. J. *Nature* **364**, 215–218 (1993).

Fruitfly origins

SIR — Begun and Aquadro¹ show that four-cutter DNA polymorphism is very different between *Drosophila melanogaster* populations from Zimbabwe and North America. They compare this result with electrophoretic data by Singh *et al.*^{2,3}. Begun and Aquadro find that there is much more similarity between populations from Benin (West Africa) and North America at the electrophoretic level than there is between populations from Zimbabwe (East Africa) and North America at the DNA level. Does this result show differences between populations in Africa or between techniques? Singh and Hale⁴ suggest that the point is technical. We disagree.

Our group, together with Michael Ashburner, put forward the hypothesis of *D. melanogaster* African origin⁵. This hypothesis does not imply that all African populations were similar, and we suggested that temperate populations recently originated from West Africa. We have studied four-cutter DNA variation in Ivory Coast (West African) flies⁶, and found much less difference with North America than Begun and Aquadro did in Zimbabwe (East African) flies. We also have compared inversion polymorphism⁷ between five populations from West Africa (including Ivory Coast and Benin) and five populations from East Africa (including Zimbabwe), and found that the two groups are both each very homogeneous, and very different from each other (see Fig. 4 in ref. 7).

Thus *D. melanogaster* populations are substantially differentiated in Africa. Unfortunately, we have only a rough idea of the degree of genetic differentiation among fruitfly populations. Present estimates are obtained from genes that are subject to selective hitch-hiking, or from allozyme and mitochondrial variation.

Françoise Lemeunier

Sylvie Aulard

*Laboratoire Populations,
Génétique et Evolution,
CNRS, 91198 Gif Cedex,
France*

Véronique Bénassi

Michel Veuille

*Institut d'Ecologie,
Université Paris VI,
7 quai St Bernard,
75005 Paris, France*

1. Begun, D. J. & Aquadro, C. F. *Nature* **365**, 548–550 (1993).
2. Singh, R. S. & Rhomberg, L. R. *Genetics* **117**, 255–271 (1987).
3. Singh, R. S., Hickey, D. L. & David, J. R. *Genetics* **101**, 235–256 (1982).
4. Singh, R. S. & Hale, L. R. *Nature* **369**, 450 (1994).
5. Lachaise, D. *et al. Evol. Biol.* **22**, 159–225 (1988).
6. Bénassi, V., Aulard, S., Mazeau, S. & Veuille, M. *Genetics* **134**, 789–799 (1993).
7. Lemeunier, F. & Aulard, S. in *Drosophila Inversion Polymorphism* (eds Krimbas C. & Powell, J. R.) 339–405 (CRC, Boca Raton, 1992).

Inversion of dorsoventral axis?

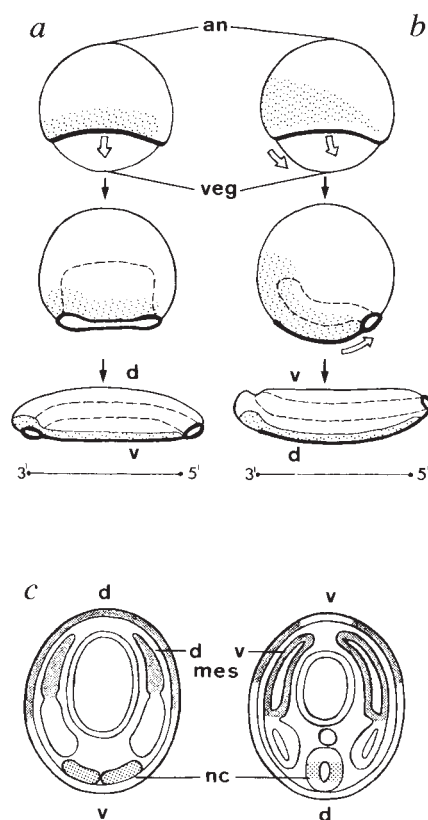
SIR — It is widely believed that the ventral nerve cord of insects and the dorsal nerve cord of vertebrates evolved independently and at opposite positions in the animal¹. Instead, we propose that the longitudinal nerve cords of insects and vertebrates derive from one and the same centralized nervous system in their common ancestor. Because the nerve cord is located ventrally in insects (classifying as *Gastroneuralia*²) and dorsally in vertebrates (*Notoneuralia*²), the ventral side of insects would then correspond to the dorsal side of vertebrates. This idea has often been discussed (beginning with Geoffroy St-Hilaire in the nineteenth century): it implies that the relative positions of the gut and nervous system, the location of the heart and the direction of the blood flow, are the same in insects and vertebrates¹.

In accordance with this notion, insects and vertebrates appear to use similar genes during formation of their differently located nervous systems. For example, in *Drosophila*³ as well as in vertebrates⁴ the *achaete-scute* complex genes are expressed in neuronal precursor cells; that is, ventrally in insects and dorsally in vertebrates, where they seem to be involved in early neurogenesis (*c* in the figure).

In addition, non-tissue-specific but homologous genes are also expressed at opposite sides in *Drosophila* and in vertebrates (outside the nerve cord anlage): The *decapentaplegic*-related genes are involved in dorsoventral patterning in both insects and vertebrates. However, the *Drosophila decapentaplegic* gene is expressed dorsally and controls the differentiation of dorsal structures, whereas its vertebrate homologue, *BMP-4*, is expressed ventrally⁵ and has ventralizing activity⁶ (*c* in the figure). Expression occurs within the ectoderm and the mesoderm, meaning that the dorsal or ventral expression is not tissue-specific or linked to any secondary cell rearrangements.

If the ventral side of insects corresponds to the dorsal side of vertebrates, how did this situation evolve? A rotation about the left-right axis (so that the dorsoventral and anteroposterior body axes are inverted, as implied in the protostome/deuterostome concept²) can be excluded because the colinear expression patterns of *Hox* cluster genes indicate that the anteroposterior body axis was fixed before the evolutionary divergence of (protostome) gastroneurians and (deuterostome) notoneurians⁷. We therefore favour an inversion only of the dorsoventral axis.

We suggest that this inversion occurred during early chordate evolution, because



Location of future nerve cord (stippled) before gastrulation (upper panels), in late gastrula (middle) and after gastrulation (lower panels) in a schematic gastroneurial animal (*a*, annelid) and a notoneurial animal (*b*, *Xenopus*). Note that after gastrulation the sides bearing the nerve cord (dorsal in chordates and ventral in annelids) both cover the now internalized vegetal hemisphere (veg) of the developing embryo. Open arrows indicate movements of blastopore margin. Anteroposterior polarity is shown as 3'–5' with reference to *Hox* cluster gene expression. Bold lines, location of the (former) blastopore; broken lines, gut anlage; an, animal; d, dorsal; v, ventral. *a*, Gastroneurial embryo showing an elongated blastopore and the ventral neurogenic region along its margins. Mouth and anus form from the anterior and posterior blastopore regions. In the developing trochophora larva of polychaete annelids the nerve cord develops along both sides of the elongated blastopore and comes to lie along the ventral side, which covers the former vegetal pole. *b*, Notoneurial embryo (*Xenopus*). During gastrulation the anterior lip of the blastopore (left) moves posteriorly around the former vegetal hemisphere while, owing to mediolateral convergence of lateral cells, neural plate and notochord elongate towards the posterior until only the anus is left open at the neural plate stage (for example, in urodeles). Note that in deuterostomes the mouth forms at the new ventral side, opposite the old gastroneurial mouth; this might have occurred as a result of adult postures similar to that of living amphioxus. *c*, Gene expression patterns in schematic cross-sections in an insect (left) and a vertebrate (right), at about the phylotypic stage⁷. Stippled region, *achaete-scute* complex genes^{3,4} in the nervous system; shaded region, non-tissue-specific *decapentaplegic*-homologues⁶; nc, nerve cord; mes, mesoderm. In *Drosophila*, *decapentaplegic* is expressed in the prospective dorsal and lateral ectoderm and in the dorsal mesoderm (mes) surrounding the embryonic midgut⁵; for better comparison the yolk is omitted and the as-yet gaping dorsal epidermis is shown closed dorsally. The vertebrate homologue of *decapentaplegic*, *BMP-4*, is expressed in the ventral mesoderm and in the ventral ectoderm (overlying the early limb buds)⁵.

the deuterostome gastrulation in chordates can be conceptually derived from the gastrulation mode seen in the presumably older (protostome) gastroneurians, such as the polychaete annelids. In polychaetes, which we assume to have more or less conserved the ancestral bilaterian mode of gastrulation, an elongated blastopore demarcates the future anteroposterior axis. Mouth and anus form at the anterior and posterior ends of this blastopore, while the nerve cord differentiates along its lateral margins². The nerve cord and the future ventral side (the mouth-bearing side oriented towards the substrate) thus come to lie opposite the animal pole, where the vegetal hemisphere was located before gastrulation (*a* in the figure).

By contrast, in deuterostome chordates only the anus derives from the blastopore. Here the initially large blastopore is being reduced to a small opening at the animal's posterior pole, because the anterior blastopore lip progresses further around the vegetal hemisphere than do the lateral and posterior lips. While the anterior blastopore lip moves posteriorly, the future dorsal structures (neural plate and dorsal mesoderm) are being expanded along the anteroposterior axis and thereby gradually overgrow, from anterior to posterior, the former vegetal side of the chordate embryo (*b* in the figure). Thus if we consider a slit-like blastopore to be ancestral also for chordates, then the dorsal midline of the early neurula — towards which ectodermal and mesodermal cells converge during gastrulation⁸ — would mark the line along which the lateral blastopore lips had fused in the chordate ancestors. This finding indicates that the chordates turned upside down early in their evolution, and henceforth were carrying the nerve cord on their dorsal side, as first proposed in 1875 (ref. 9).

The proposed homology of nerve cords and the assumed inversion of the dorsoventral body axis in early chordates need not change the current view about genealogical relationships between the animal groups now considered 'gastroneurians' and 'notoneurians'. But it may change our view of what the common ancestor of insects and vertebrates could have looked like.

D. Arendt

K. Nübler-Jung

Institut für Biologie I (Zoologie),
Albert-Ludwigs-Universität Freiburg,
79104 Freiburg, Germany

1. Romer, A. S. & Parsons, T. S. *The Vertebrate Body* (Saunders, Philadelphia, 1977).
2. Fioroni, P. *Rev. suisse Zool.* **87**, 261–272 (1980).
3. Martin-Bermudo, M. D. *et al. Development* **113**, 445–454 (1991).
4. Zimmerman, K. *et al. Development* **119**, 221 (1993).
5. Jones, C. M. *et al. Development* **111**, 531–542 (1991).
6. Jones, C. M. *et al. Development* **115**, 639–647 (1992).
7. Slack, J. M. W. *et al. Nature* **361**, 490–492 (1993).
8. Keller, R. *et al. Dev Dyn.* **193**, 199–217 (1992).
9. Nübler-Jung, K. & Arendt, D. *Roux Arch. dev. Biol.* **203**, 356 (1994).

Evolution in real time

Jeremy J. D. Greenwood

The Beak of the Finch. By Jonathan Weiner. Knopf, New York/Jonathan Cape, London: 1994. Pp. 332. \$25, £18.99.

DURING his visit to the Galápagos Islands in 1835, Charles Darwin made a collection of finches. But although realizing that they comprised several species, he was so little impressed with the differences between the birds of different islands that he failed to record accurately from which island some of his specimens came. It took John Gould, working on Darwin's ornithological collections, to point out that this was a unique group of birds, closely related to each other but with great variation in bill shape. Of course, once Gould had done so, it was Darwin who made the intellectual leap, remarking in his published journal of the *Beagle* voyage:

Seeing this gradation and diversity of structure in one small, intimately related group of birds, one might really fancy that from an original paucity of birds in this archipelago, one species had been taken and modified for different ends.

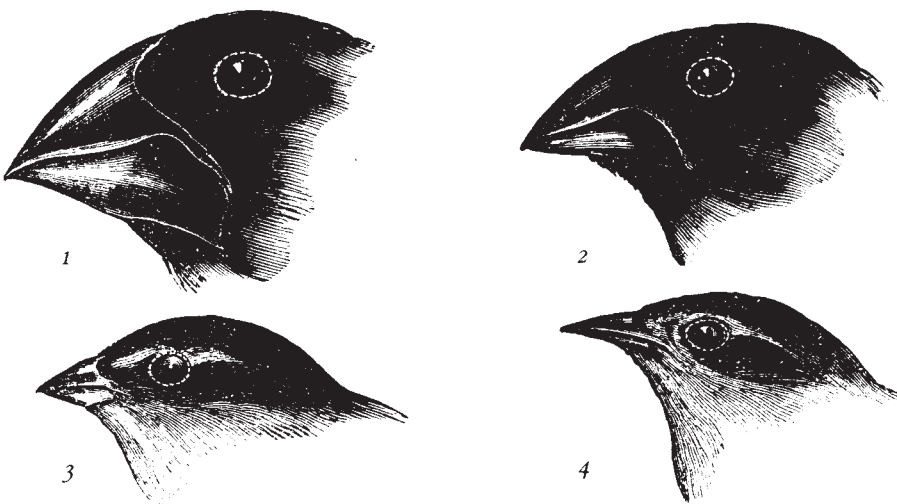
The beaks of the birds have indeed become particularly "modified for different ends" — hence Jonathan Weiner's title and his reproduction of the illustration of the heads of four of the species from Darwin's journal. His book is not however about Darwin's own very limited work on the finches but about that of Peter and Rosemary Grant and their co-workers more than a century later. It is a story not just of a major research project but also of how the work fits into the broader realms of evolutionary biology.

The Grants first visited the Galápagos archipelago in 1973, having concluded that its finches, showing much variation within a closely related group and living in largely undisturbed localities, were ideal subjects for exploring the importance of competition in the evolution of avian species. They have been returning year after year, with their children, Thalia and Nicola, often accompanying and assisting them; Thalia's illustrations are a striking addition to Weiner's text.

Many other zoologists have visited the Galápagos Islands since Darwin, filling the world's museums with specimens of biota, including finches. None made such a great contribution as the late David Lack, who stayed there for a full five months during 1938 and 1939, studying living finches both in captivity and in the wild, observing how they live. Noting that different species could commonly be seen feeding together on the same seeds, he concluded that, although generic differences were adaptively related to the birds'

food, specific differences were not. But by the time his paper on these findings was published (it was delayed by the Second World War), he had changed his mind: species differences were indeed adaptive, he argued, with competition between species moulding the differences.

Lack's initial misfortune was to have studied the birds during a wet season,



Four of Darwin's finches, from his *Journal of Researches*: (1) large ground finch; (2) medium ground finch; (3) small tree finch; (4) warbler finch.

when food was plentiful. The Grants have gone much further by extending their work to include not only the variation in conditions during the course of a year but the dramatic variation from year to year, particularly those brought about by El Niño, an extensive prolonged warming of the eastern tropical Pacific. They have watched the finches in bad times as well as good, observing how natural selection and competition vary in strength over time; indeed, how natural selection may vary in direction as conditions change. They have studied whole populations of birds at the individual level, recording the fortunes of each bird and its descendants, thereby building up comprehensive genealogies. They have measured beaks and many other characteristics, physical and behavioural. They have observed how food supplies vary not just in terms of absolute quantity but also in the relative abundance of different types of seed. They have measured how easy it is to remove husks from the different seeds and related this to the birds' beaks.

Such long-term fieldwork has not been fashionable. It has been more popular to grind up tissues in the laboratory to study

the variation of molecules and to conclude, on the basis of agreement between the observations and mathematical theory, that most evolutionary changes are selectively neutral. Because the functional significance of the variants studied has usually been both obscure and difficult to measure with great precision, the conclusion was scarcely surprising. This laboratory work has perhaps been attractive because it can be done quickly, within the span of a single funding grant, and it does not require the deep knowledge of natural history that can be acquired only through years of experience. It seems glamorous too: compare its complex apparatus and white coats with the notebooks, binoculars and work-stained gar-

ments of the island naturalist.

Yet the Grants have the crucial pieces of equipment, the eyes and brain of the thoughtful naturalist. As a result, it is they who have added brick after brick to the growing edifice of evolutionary biology. Their greatest triumph has been to show not only that natural selection can be measured in the field, so that providing one has also measured heritability in the field, one can predict the outcome of selection, but also that such predictions can be precisely correct.

Weiner's book is not an exposition of evolutionary theory. But by telling a lively story of one group's contribution, it conveys important elements of the theory. It also shows that science is an exciting endeavour done by three-dimensional people, not the dull abstraction studied by automata — the image that turns off so many young people. The style and presentation of this book are first class. It is one of the best pieces of science writing that I have read in a long while. □

Jeremy J. D. Greenwood is director of the British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU, UK.

From *The Beak of the Finch*

Fancying the fancy

James L. Gould

Sexual Selection. By Malte Andersson. Princeton University Press: 1994. Pp. 599. \$65, £52.50 (hbk); \$24.95, £19.95 (pbk).

DARWIN distinguished between natural and sexual selection: the former relates to survival and is similar for both sexes; the latter centres on obtaining mates and reproducing, and often differs between males and females. Darwin then subdivided sexual selection into male contests and female choice. The entire concept of sexual selection was controversial for decades, and female choice was so ridiculed and ignored that the first clear demonstrations of it in action (in widowbirds and guppies) appeared only a dozen years ago. Now the author of the original widowbird study, Malte Andersson, has produced his long-awaited monograph on sexual selection.

Andersson's book is a testimony to the

explosive growth of interest in sexual selection. Its impressive list of 2,300 citations, for instance, has a median date of only 1985. And for his survey to be manageable, he has had to omit so many topics that there is only about a 50 per cent overlap of references with another recent book on the subject.

What is missing? The whole question of the evolution and maintenance of sex and the two-sex system is purposely left out. Too bad, but probably necessary. Sexual selection in humans is omitted, although on the mistaken premise that there is no work on this fascinating and well-studied subject. Almost everything relating to physiological mechanisms is missing, probably because behavioural ecologists have traditionally had little interest in the 'plumbing'.

What is slighted? Ironically, Andersson credits Darwin with exaggerating the importance of female choice at the expense of male contests; in fact, Darwin, like Andersson, simply found female choice more interesting, and it receives an appropriately disproportionate amount of attention here — and why not, since the question of why females should be programmed to prefer style over substance is

one of the most intriguing in all of biology. The sensory-bias hypothesis gets relatively little discussion, but then this idea, well known to ethologists for some time, has only recently begun to gain ground among behavioural ecologists; Andersson does point out that evidence for sensory bias is beginning to accumulate rapidly. Species recognition and frequency-dependent selection also get minimal treatment.

What is emphasized? In addition to female choice, Andersson devotes much space to an analysis of the models of sexual selection, particularly of female choice. This useful focus provides a good description of the general predictions of the models, although there is no help with the mathematical underpinnings or the formal logic. This attention apparently flows from Andersson's opinion, repeatedly stated here, that detailed mathematical models drive experimentation. This may be true for Andersson, but I doubt it characterizes many of the rest of us — and just as well, since the modellers have often failed to anticipate the experimental findings of researchers muddling along on intuition alone. Andersson also devotes a surprising and welcome amount of attention to sexual selection in plants, although again the absence of physiological thinking limits the analysis.

In general, Andersson provides a very balanced discussion of each of the issues he chooses to consider, with due consideration for minority views and possible objections. His healthy scepticism in the face of the various intellectual fads that have come and gone in this exciting field is epitomized by his question: "Is this because reality is becoming more accurately portrayed, or is there a new myth in the making?" He does an excellent job in isolating the important problems and worrisome loose ends.

Sexual Selection is more a reference than a text; it is not really suitable for a course because Andersson has, quite reasonably, chosen breadth over depth. So most classic work — the alternative mating strategies of bluegill sunfish, for instance — gets a paragraph at most; less memorable tests usually evoke only a sentence. This breathtaking pace and need for abbreviation is reflected in the remarkable paucity of illustrations — 120 in 600 pages. About two dozen studies, however, are highlighted in modest depth, and remind us how interesting the details really are. The book will no doubt serve as the major inspiration for many subsequent experiments, and thus will contribute, as Andersson must hope, to its own rapid obsolescence. □

James L. Gould is in the Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08540-1003, USA.



THE great stūpa (Buddhist monument) at Sāncī, India, was built in brick by the Mauryan king Asoka (c. 268–232 BC) and enlarged and encased in stone by the Śuṅga kings around 100 BC. The sculptural decoration is concentrated on the gateways, added in the Christian era. The north gate, shown here, depicts the Buddha's 'wheel of law'. From *The Penguin Encyclopedia of Classical Civilizations* ed. A. Cotterell (Viking, \$29.95).

Insects and vortices

R. McNeill Alexander

The Evolution of Insect Flight. By A. K. Brodsky. Oxford University Press: 1994. Pp. 229. £55, \$82.50.

THE mechanics of insect wing movements can be understood only in the light of the complex anatomy of the thorax and wing bases. The mechanisms of flight present formidable problems in aerodynamics. Happily, Andrei Brodsky is expert in both. He has been doing fine research on insect flight for 25 years but, published in Russian, it was little known in the West



Flying fritillary — taken from John Brackenbury's *Insects in Flight* (Blandford, 1992; £18.99). His latest book *Insects: Life Cycles and the Seasons* (Blandford, £19.99), to be published later this month, once again provides a stunning showcase for the author's unique photographic techniques.

until a paper by him appeared in the *Journal of Experimental Biology* in 1991. There he used spectacular photographs of smoke flowing past the wings of a flying butterfly to reveal the vortices on which flight depends. We have been familiar for several years with the vortex patterns of bird flight; vortex rings produced in the downstroke only at low speeds, and wavy trailing vortices in fast flight. Brodsky has shown that insect flight depends on a different pattern, a chain of linked vortices, one for each up- or downstroke. This results from the circulation of air around the wings being reversed between

each stroke and the next, a pattern apparently used by hummingbirds but no other birds.

The book starts with a superb general account of the structure of insect wings and thoraxes, of the mechanism of wing beating and of the kinematics and aerodynamics of flight. It is illustrated by diagrams that time and again gave me a shock of pleasure as I came to understand points much more clearly than I had done before. Other illustrations are complicated drawings that cannot be appreciated without careful study.

The general account of flight occupies a third of the book, and will be widely useful. The rest, on the evolution of flight, will be valuable to specialists who cannot read Brodsky's papers in their original Russian, but most readers will be content to skip the details and extract just the broad picture. Brodsky uses fossil evidence, comparative anatomy and studies of living insects to identify three main lineages of insect flight and nine types of kinematics.

There is no physiology in this book, but for the anatomy, mechanisms and evolution of insect flight it has no rivals. □

R. McNeill Alexander is in the Department of Pure and Applied Biology, University of Leeds, Leeds LS2 9JT, UK.

Deep questions, flawed answers

N. David Mermin

Observational Foundations of Physics. By Alan Cook. Cambridge University Press: 1994. Pp. 164. £27.95, \$39.95 (hbk); £12.95, \$18.95 (pbk).

THE second is now defined as the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium-133 atom; since 1991 the metre has been defined as 1/299,792,458 of the distance travelled by light in a second. What implications do these definitions have for the form and content of physical theory?

According to Alan Cook the implications are profound. The first definition requires the time evolution of a quantum mechanical state to be given by a first-order differential equation, and together they lead to the invariance of the relativistic interval along with the accompanying geometry of four-dimensional vectors. Cook also argues that there is a link between defining time by a classical periodic process and equations of motion of the Hamiltonian form. In all these cases, he suggests, much of the detailed

formal structure of theoretical physics may be determined by how we locate things in space and time.

Many other deep questions are examined in this little book. What does chaos in classical dynamical systems do to the notion that the aim of science is to make successful predictions? How should it lead us to reexamine the foundations of statistical mechanics? What, for that matter, is the nature of probability? Is it subjective or objective? Does it deal with frequencies or degrees of belief? Why, when you come to think of it, is mathematics so extraordinarily good at accounting for physical behaviour?

The questions are stimulating and Cook offers some highly original answers. Particularly intriguing (and provocative) is his notion that the remarkable power of mathematics can be traced to the fact that physical systems themselves might be viewed as concrete realizations of the basis vectors for the irreducible representations of groups.

But more often than not I found his answers unpersuasive. The questions are deep and difficult, and a book so short and comprehensive can do little more than point to possible directions. The arguments sketched by Cook, however, often suffer from a lack of clear logical structure — a failure to distinguish between what follows from the premises and what is simply put in by hand.

Furthermore, when an assertion is clear and precise, it is often simply incorrect. Cook states, for example, that a symmetry group describing reflection in a single plane has three members (it has two); he asserts that the number of operations in a space group is finite (it is infinite); he takes as fundamental that it is impossible to observe distant events except with electromagnetic radiation (what about neutrino signals, and why are millions being spent to build gravitational wave detectors?); he declares that a quantum mechanical Hamiltonian must be Hermitian because "all wave functions are orthonormal" (an absurdity — it is because the time evolution must preserve orthonormality); two pages after a statement that the equations of classical dynamics are generally not of a certain form, we are told how they can always be put in just that form. The accumulation of such small infelicities undermines confidence in the broader and more obscure parts of the argument.

This, in short, is a book that takes up many interesting questions but offers answers that are fragmentary and flawed. If it stimulates others to do a better job, it will have served a useful purpose. □

N. David Mermin is in the Laboratory of Atomic and Solid State Physics, Cornell University, Clark Hall, Ithaca, New York 14853-2501, USA.



THE Malaspina Glacier, Alaska, is the largest of the world's piedmont glaciers and a superb example of folding due to glacial flow. This picture comes from the new paperback edition of *Glaciers* by Michael Hambrey and Jürg Alean (Cambridge University Press, £12.95, \$15.95), a well-illustrated introduction for lay people. Hambrey has also just written a more detailed textbook on the science of regions dominated by snow and ice: *Glacial Environments* (UCL Press, £40 (hbk), £14.95 (pbk)).

The map book

David W. Hughes

The Guide to the Galaxy. By Nigel Henbest and Heather Couper. Cambridge University Press: 1994. Pp. 265. £35, \$49.95 (hbk); £17.95, \$24.95 (pbk).

IN the 1930s the problem of surveying our own Galaxy was likened to the problem of drawing a map of London based on observations made when sitting on the back of one of the lions at the foot of Nelson's Column on a foggy day. It was the cosmic equivalent of the forest problem: one could not see the wood for the trees.

But things have changed enormously in the intervening years. Instead of just relying on the light from stars, astronomers now have information from radio waves, millimetre waves, infrared radiation, ultraviolet radiation, X-rays and gamma rays. What was once supposedly an assemblage of ageless stars gyrating sedately through empty space is now a maelstrom of molecular and atomic gases, interspersed with clouds of dust and intertwining magnetic fields, all this being churned by the birth throes and death throes of stars.

Nigel Henbest and Heather Couper lay special emphasis on the pan-spectral nature of today's attempts to discern the shape and structure of the Galaxy we live in. They also stress the way in which our

understanding has developed, and their book abounds with portraits of past astronomers who have dedicated their lives to galactic mapping.

The Guide to the Galaxy is aimed at the nontechnical reader. There are no equations and very few line drawings. What the book does have, however, is both a superb collection of images and maps and an eminently readable text. The maps are especially fine and have been drawn for the book by Julian Baum. Here we see diffuse atomic hydrogen regions and hydrogen molecular clouds represented by different colours and intensities, these regions being interspersed by nebulae, stars and star associations. These maps make the whole structure of the Galaxy spring to life and the book's careful positioning of maps and images is a great aid to understanding the complexity of our star system.

The text is also first-class. After exploring the local neighbourhood and being convinced that our glade is just a typical neck of the Galactic wood, we are then encouraged to wander along the Perseus, Orion and Sagittarius arms before being transported to the mysterious regions of the Galactic Centre. The bubbling enthusiasm of the two authors springs from every page. They are Galaxy fans and gleefully enjoy taking the reader on this voyage of exploration. □

David W. Hughes is in the Department of Physics, University of Sheffield, The Hicks Building, Sheffield S3 7RH, UK.

New in paperback

The Private Lives of Albert Einstein by Roger Highfield and Paul Carter. Faber and Faber, £8.99. The authors "have used Einstein's personal correspondence, some of it discovered as recently as 1986, to produce a narrative that reveals him to have experienced angst and joys not uncommon to most mortals" (Arthur I. Miller, *Nature* **366**, 371; 1993).

Toward the Habit of Truth: A Life in Science by Mahlon Hoagland. The autobiographical memoirs of one of the pioneers of the biochemistry of protein synthesis. Norton, \$12.95, £9.95. Reviewed by Sydney Brenner in *Nature* **345**, 675; 1990.

From Stone to Star: A View of Modern Geology by Claude Allègre. Harvard University Press, \$16.95, £13.50. "Highly readable and deeply informative. . . An excellent introduction to the exciting geohistorical discoveries made in recent decades. . . it will serve alike the needs of the geologist and the general reader" (Gordon L. Herries Davies, *Nature* **358**, 719; 1992).

Fossil Horses: Systematics, Paleobiology, and Evolution of the Family Equidae by Bruce J. MacFadden. Cambridge University Press, £15.95, \$29.95. "A timely and readable text, a good advertisement for the biological fruits that the palaeontological tree can bear" (Adrian Lister, *Nature* **365**, 118; 1993).

DNA Fingerprinting by M. Krawczak and J. Schmidtke. BIOS Scientific, £16, \$30. An introduction to genetic fingerprinting and profiling. Aimed at lawyers and students of medical genetics, forensic science or medicine.

The Hutchinson Dictionary of Science edited by Paul Lafferty and Julian Rowe. Helicon, £12.99. Contains over 5,000 A-Z entries and 20 "feature articles".

How to Write about Biology by Jan Pechenik and Bernard Lamb. HarperCollins, £6.99. A guide for students that covers, among other topics, lab reports, essays, research papers and proposals, job and graduate-study applications and oral presentations. Geared towards British and European undergraduates and postgraduates, the book is an adaptation of the second edition of Pechenik's US forerunner, *A Short Guide to Writing about Biology*.

A Stillness in the Pines: The Ecology of the Red-Cockaded Woodpecker by Robert W. McFarlane. Norton, £10.95. An ornithologist looks at the plight of this endangered species that is unique to the southeastern United States.

In vitro evolution of new ribozymes with polynucleotide kinase activity

Jon R. Lorsch^{*†} & Jack W. Szostak^{*‡}

^{*} Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

[†] Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

[‡] Department of Genetics, Harvard Medical School, Longwood Avenue, Boston, Massachusetts 02115, USA

We have isolated a large number of polynucleotide kinase ribozymes from a pool of RNA molecules consisting of an ATP-binding domain flanked by regions of random sequence. Different classes of kinases catalyse the transfer of the γ -thiophosphate of ATP- γ S to the 5'-hydroxyl or to internal 2'-hydroxyls. An engineered version of one class is able to catalyse the transfer of thiophosphate from ATP- γ S to the 5'-hydroxyl of an exogenous oligoribonucleotide substrate with multiple turnover, thus acting as a true enzyme.

SELECTION and evolution *in vitro* have allowed the isolation of nucleic acids (aptamers) that bind a variety of small molecules¹⁻⁸ and proteins⁹⁻¹². These techniques have also been used to alter the activities and functions of ribozymes¹³⁻¹⁵. A number of RNA molecules that catalyse self-modifying reactions have been isolated from libraries of partially randomized tRNA¹⁶ and random sequence RNAs¹⁷. RNA molecules have also been generated that catalyse the interconversion of two hindered conformational isomers¹⁸. We wanted to use a pre-existing RNA aptamer as the basis for a selection for ribozymes that use a biologically relevant small molecule substrate. ATP is arguably the most important biological cofactor. It is used in many metabolic reactions, for regulatory purposes, and by enzymes that replicate and maintain the genetic material. The polynucleotide kinases are one such class of enzymes. They are widely distributed in nature and are believed to be involved in a number of important biological processes including DNA damage repair, and RNA splicing and ligation^{19,20}. If the 'RNA world' evolved to the point of containing complex cells, ribozyme kinases capable of fulfilling these functions would have been required. An indication that this might have been possible comes from the demonstration that the *Tetrahymena* group I intron can transfer a 3'-phosphate from one oligoribonucleotide to the 3'-hydroxyl of another²¹. We set out to use the ATP aptamer domain⁴ isolated previously in our laboratory as the starting point for a selection for polynucleotide kinase ribozymes.

Selection

We designed a pool of random sequence RNAs, using the minimal ATP aptamer as a core structure⁴. By creating a pool that was predisposed to bind ATP specifically and with high affinity we hoped to increase the likelihood of generating molecules with ATP-dependent kinase activity. The ATP aptamer core was surrounded by three regions of random sequence of 40, 30 and 30 nucleotides in length, respectively (Fig. 1). The ATP-binding domain itself was mutagenized such that each base had a 15% chance of being non-wild-type, to allow for changes in the aptamer sequence that might be required for optimal activity. To increase the likelihood of finding active molecules, we attempted to create a pool containing as many different molecules as possible. Because it is difficult to obtain an acceptable yield from the synthesis of a single oligonucleotide of this length (174 nucleotides), we made two smaller DNA templates and linked them together to generate the full-length DNA pool (Fig. 1)¹⁷. Transcription of this DNA yielded between 5×10^{15} and 2×10^{16} different RNA molecules.

To select for catalytic activity it is necessary to tag active molecules so that they can be separated from inactive ones. We

incubated the random sequence RNA pool with ATP- γ S (Fig. 1) and selected for transfer of the thiophosphate from ATP- γ S to the RNA by chromatography on a thiopyridine-activated thiopropyl Sepharose column, which forms disulphide bonds with RNAs containing thiophosphate groups. Molecules without thiophosphates could then be washed away under denaturing conditions. RNAs linked via a disulphide to the column matrix were eluted with an excess of 2-mercaptoethanol. Reverse transcription, PCR and transcription yielded a new RNA pool enriched in active molecules. This process comprised one cycle of selection.

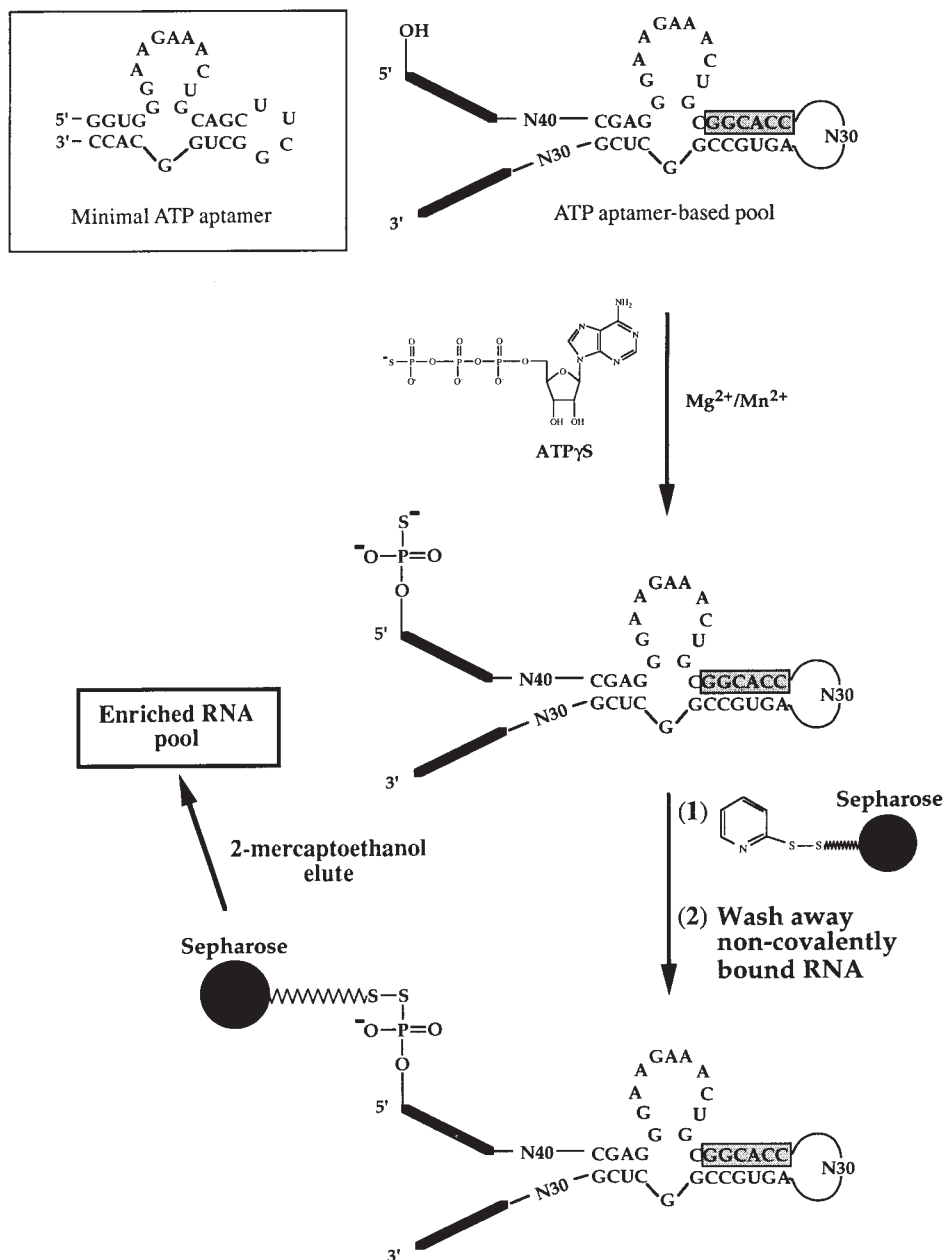
Before each cycle of the selection, the pool RNA generated by transcription was exhaustively dephosphorylated with calf intestinal alkaline phosphatase to remove the 5'-triphosphate, and any other phosphates that might have been transferred to the RNA by autophosphorylation during transcription.

Our selection demanded only that an RNA molecule contain a thiophosphate in order for it to be isolated. Reactions that could have been selected for include: transfer of the γ -thiophosphate from ATP- γ S to the 5'-hydroxyl of the RNA (analogous to the reaction catalysed by T4 polynucleotide kinase), to the 3'-end of the RNA, to an internal 2'-hydroxyl, or even to a group on one of the bases. Transfer of a diphosphate (or perhaps the entire triphosphate) instead of a single thiophosphate is also possible for all of these reactions. A splicing reaction, in which ATP- γ S displaces one of the first few nucleotides of the RNA in a manner analogous to the reaction catalysed by the group I introns, could also occur. However, cleavage of more than the first few bases of the RNA would result in a molecule lacking a 5'-primer binding site, and such a molecule would not be amplified during the PCR step of the selection. Similarly, any reaction that blocked reverse transcription would not be selected for.

Figure 2 shows the progress of the selection in terms of the fraction of the pool RNA that bound to the thiopropyl Sepharose and was eluted with 2-mercaptoethanol. Initially, about 0.5% of the RNA bound nonspecifically to the matrix and was eluted with 2-mercaptoethanol. After five cycles of selection, greater than 20% of the pool RNA reacted with thiopropyl Sepharose. As we estimated that there were at least 10,000 different molecules left in the pool at this stage, we chose to increase the stringency of the selection in the succeeding cycles by lowering the ATP- γ S concentration and the incubation time, in order to try to isolate the most active catalysts. Because our selection sampled sequence space very sparsely (there are $4^{100} \approx 10^{60}$ possible 100-mers, but only $\sim 10^{16}$ different molecules in our pool), active molecules are likely to be sub-optimal catalysts. We therefore chose to perform three cycles of mutagenic

FIG. 1 Pool design and selection scheme for kinase ribozymes. Top left: Minimal ATP aptamer⁴. Top right: random RNA pool built around aptamer structure. The pool contained three regions of random sequence (N) for a total of 100 randomized bases. The aptamer region was mutagenized to a level of 15%. The *BanI* site used to ligate the two halves of the pool is shown in grey. Constant primer binding sites are shown as thick lines. Pool RNA was allowed to react with ATP- γ S, and thiophosphorylated molecules were isolated by reaction with thiopyridine-activated thiopropyl sepharose. Nonspecifically bound molecules were removed by washing under denaturing conditions, and active molecules eluted with 2-mercaptoethanol. Constant regions: 5'-GGAACCUCUAGGUCAUUAAGA-3' (5'-end); 5'-ACGUCAGAAGGAUCCAAG-3' (3'-end).

METHODS. The random pool was constructed as previously described (ref. 17). The pool was incubated at 23 °C with ATP- γ S under conditions designed to promote the formation of RNA tertiary structure (400 mM KCl, 50 mM MgCl₂, 5 mM MnCl₂, 25 mM HEPES, pH 7.4). Mn²⁺ was included because of its ability to coordinate phosphorothioates³⁸. Streptavidin agarose immobilization of pool RNA was used during the first seven cycles to prevent pool aggregation¹⁷. After cycle 7, the ATP- γ S reaction step was performed in solution (1 μ M RNA). For the first cycle, 2.4 mg (40 nmol; 5 pool equivalents) of random pool RNA was used, in the second cycle 150 μ g (2.4 nmol) RNA was used, and in succeeding cycles 60 μ g (1 nmol) was used. The selection step was performed by incubating the RNA with thiopyridine-activated thiopropyl Sepharose-6B (Pharmacia) in 1 mM EDTA, 25 mM HEPES, pH 7.4 for 30 min at room temperature. The resin was then washed with 20 column volumes each of wash buffer (1 M NaCl, 5 mM EDTA, 25 mM HEPES, pH 7.4), water, and finally 3 M urea, 5 mM EDTA. RNAs linked to the resin via a disulphide were eluted with 0.1 M 2-mercaptoethanol in 0.5 $\times$ wash buffer.



PCR^{17,22} to allow the evolution of improvements in the active molecules. The combined effect of increasing the stringency and performing mutagenic PCR was to increase the activity of the pool by nearly three orders of magnitude from cycle 6 to cycle 13 (Fig. 2*b* and *c*).

Composition of the final pool

After 13 cycles of selection RNA molecules from the pool were cloned and sequenced. The 50 clones sequenced (Fig. 3*a*) fall into seven classes of two or more closely related molecules (19 clones) and 31 unique sequences. Each class of sequences represents molecules with a common ancestor that acquired mutations during the course of the mutagenic PCR done in cycles 7–9 of the selection. We focused our analysis on the seven major classes of molecules.

Comparison of the sequences reveals significant conservation of the sequence of the original ATP binding site in some of the active RNAs. Figure 3*b* shows the putative structures for the ATP aptamer regions from classes I, III, IV and V, the classes for which an aptamer-like structure can be drawn. It appears

that classes I and III have changed significantly from the original ATP binding domain, whereas classes IV and V are only slightly different from the ATP aptamer consensus sequence⁴. Either the right or left hand stems of the class II, VI and VII aptamer regions appear to be missing, and it seems likely that these molecules have found novel modes of binding their substrates. Using run-off transcription of synthetic DNA oligonucleotides²³ we made the RNAs corresponding to the class I, III, IV and V aptamer regions and found that only the class IV aptamer RNA binds detectably to C-8 linked ATP agarose⁴, suggesting that the corresponding classes of kinases bind ATP- γ S more weakly or in a different way than the original ATP aptamer.

The catalysed reactions

We tested pool 13 RNA, and the members of each of the major classes of kinases to determine what reactions they catalyse. PEI cellulose thin-layer chromatography (TLC)²⁴ of nuclease P1 digests of autothiophosphorylated RNA shows two major radiolabelled products (Fig. 4*a*), demonstrating that at least two different reactions are catalysed by the pool 13 RNAs. If a par-

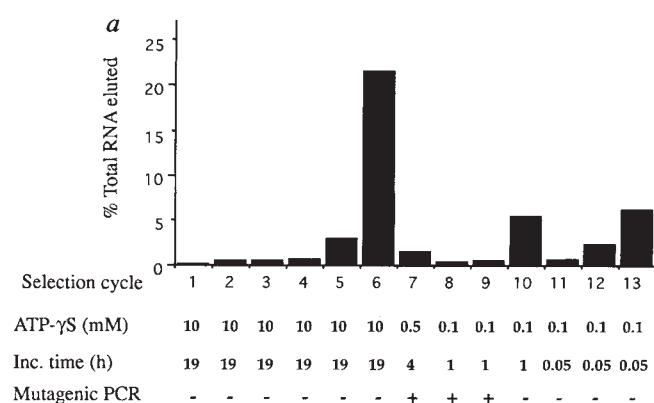
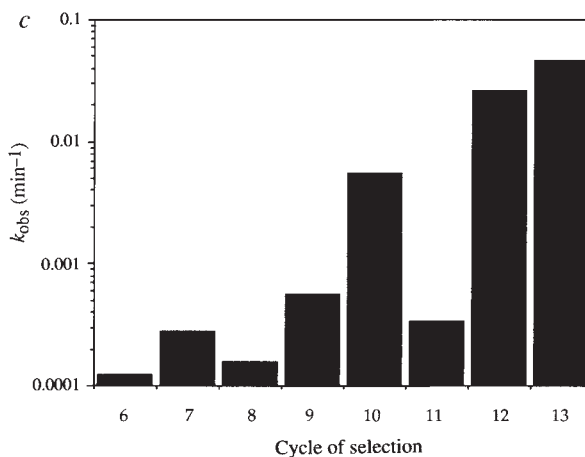
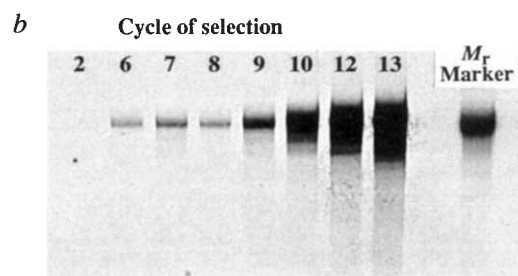


FIG. 2 Evolution of new ribozymes. *a*, Summary of kinase selection. Per cent RNA eluted by 2-mercaptoethanol from the thiopyridine-activated thiopropyl Sepharose at each cycle of the selection. Background sticking and elution of RNA from the resin is $\sim 0.5\%$. Cycle 1 background is lower because of the higher RNA/resin ratio. Conditions for each ATP- γ S reaction step are shown at bottom. Mutagenic PCR was performed before cycles 7, 8 and 9 in order to allow the evolution of beneficial base changes. *b*, Activity of the RNA pool at various cycles of selection. RNA was allowed to react with trace ATP- γ - 35 S for 3 h under selection buffer conditions. The reactions were then quenched and run on denaturing 10% polyacrylamide gels. Labelled pool RNA was run as a marker. *c*, k_{obs} of pool RNA for selection cycles 6–13. Reactions were performed with 100 μ M ATP- γ S, and a time point was chosen such that $<20\%$ of the active RNA had reacted. At cycle 6 the activity of the pool could be readily detected. The following seven cycles increased the activity by nearly three orders of magnitude. The drop in k_{obs} in cycle 8 is presumably due to the effects of mutagenic PCR, coupled with the fact that the pool was no longer immobilized on streptavidin agarose in this cycle. Cycle 11 activity declined for unknown reasons. METHODS. Mutagenic PCR was done as described^{17,22}, except that only 30 total cycles of PCR were done at each round to yield $\sim 2\%$ mutagenesis. Reactions of pool RNAs (1 μ M) were performed at 23 °C either with trace ATP- γ - 35 S (*b*), or with 100 μ M ATP- γ S plus additional trace

ticular RNA molecule transfers the γ -thiophosphate from ATP- γ S to its own 5'-hydroxyl, the nuclease P1 digestion should yield labelled GMP- α S, as all of the RNAs begin with guanosine. All members of classes I, II, III, V and VI yield GMP- α S as the sole nuclease P1 digestion product (Fig. 4b), indicating that they are 5'-kinases. Classes IV and VII, on the other hand, yield a nuclease P1 digestion product that does not migrate from the origin in this TLC system. Both RNase T2, which hydrolyses RNA to nucleotide 3'-monophosphates, and nuclease P1 digestion of reacted class IV and class VII RNAs, give products that run as molecules with charges of -5 to -6 on DEAE-cellulose TLC plates, using a solvent system that separates based upon the charge of the RNA fragment^{25,26}. These data are consistent with class IV and VII RNAs being internal 2'-kinases, because neither nuclease P1 nor RNase T2 can cleave at 2'-phosphorylated sites^{24,26}. The products of these digestions, then, should be 35 S-labelled dinucleotides with 5'-phosphates or 3'-phosphates (for nuclease P1 and RNase T2 digestions respectively) and 2'-mono- or diphosphates.

Experiments in which the RNAs were allowed to react with unlabelled ATP- γ S, and were then purified and reacted with [γ - 32 P]ATP and T4 polynucleotide kinase support the proposal that classes I, II, III, V and VI are 5'-kinases, and that classes IV and VII phosphorylate some internal site. As expected, reaction products from classes I, II, III, V and VI cannot be labelled by T4 polynucleotide kinase, consistent with their being 5'-kinases. Class IV and VII RNAs, on the other hand, are efficiently labelled by T4 polynucleotide kinase after they have been allowed to react with ATP- γ S. Furthermore, this labelled RNA can be purified on a thiopyridine-activated thiopropyl Sepharose column, demonstrating that the thiophosphate label is not lost



ATP- γ - 35 S (*c*). 10 mM dithiothreitol (DTT) was included in the reactions. Reactions were quenched by the addition of one volume of 150 mM EDTA, 20 mM DTT in 95% formamide. Reactions were analysed by electrophoresis on 10% polyacrylamide/8 M urea gels. Quantitation was performed using a Phosphorimager (Molecular Dynamics). A known amount of ATP- γ - 35 S was spotted on the gels as a standard.

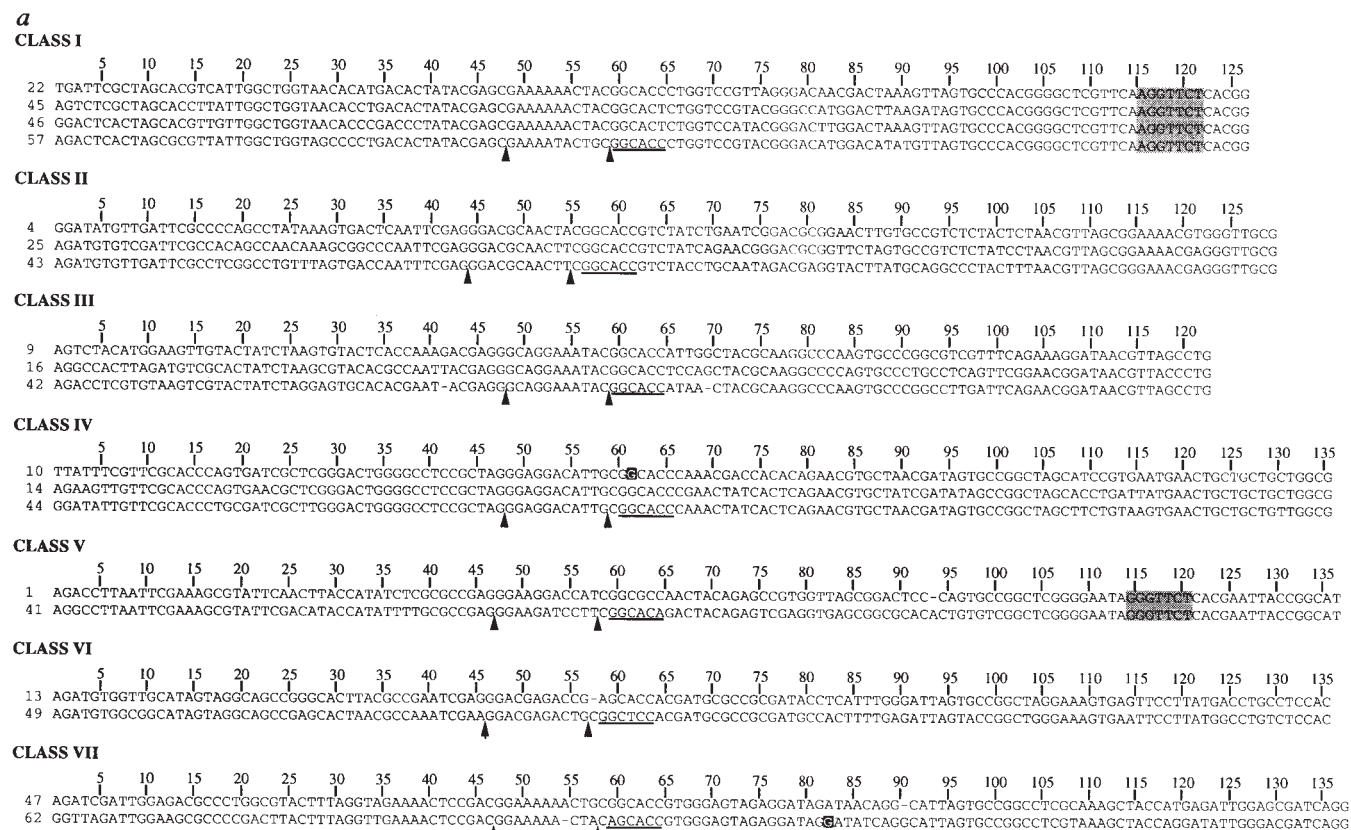
during the reaction with ATP and T4 polynucleotide kinase. Thus, the class IV and VII kinases do not catalyse reactions involving their 5'-hydroxyls.

Further evidence for the 2'-kinase hypothesis is provided by partial alkaline hydrolysis of the autothiophosphorylated, 5'- 32 P-labelled RNA (Fig. 5). For RNAs from both classes IV and VII, a gap is seen in the hydrolysis ladder of the autothiophosphorylated material that is not present in the ladder made with unreacted RNA. The missing bands can be most easily explained if the 2'-hydroxyls at these positions are thiophosphorylated, thus preventing base-catalysed RNA hydrolysis. This experiment has also allowed us to localize the position of thiophosphorylation: G62 in Kin.10 (class IV) and G83 in Kin.62 (class VII). G62 is in a putative helix within the ATP aptamer region of Kin.10 and G83 is in the random loop between the two halves of Kin.62's aptamer domain.

This result also suggests that M-MLV (Moloney murine leukemia virus) reverse transcriptase is able to read through 2'-phosphates on template RNAs.

Kinetic analysis of kinase ribozymes

Kinetic analysis of the most active clone from each of the four major classes of kinases has revealed that they all obey the standard Michaelis-Menten kinetics expected of molecules possessing saturable substrate binding sites. Values of k_{cat} and K_m are shown in Fig. 3, and range between 0.03 and 0.37 min⁻¹ and between 41 and 456 μ M, respectively. The k_{cat} for class I-IV ribozymes compares favourably with corresponding values for naturally occurring ribozymes, which range from 0.04 to 2 min⁻¹ (refs 27–32). Comparison of k_{cat}/K_m is difficult because most natural ribozymes have oligonucleotide substrates that

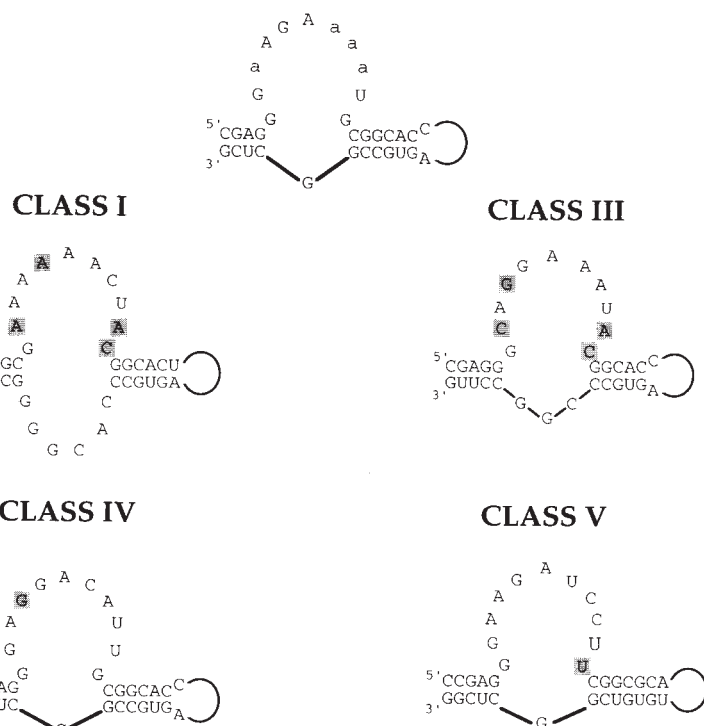


c

Kinase class (clone)	k_{cat} (min^{-1})	K_m (μM)
Class I (Kin.46)	0.37 ± 0.01	456 ± 57
	0.23 ± 0.02	116 ± 41
	0.36 ± 0.02	352 ± 85
Class II (Kin.25)	0.20 ± 0.02	41 ± 15
	0.33 ± 0.02	42 ± 11
Class III (Kin.42)	0.07 ± 0.005	50 ± 13
	0.10 ± 0.016	58 ± 28
Class IV (Kin.44)	0.03 ± 0.001	276 ± 25
	0.03 ± 0.001	200 ± 22

FIG. 3 Major kinase classes. **a**, Sequences of the seven classes of kinases (50 clones sequenced). Arrows delimit the ATP aptamer conserved loop. The *BanI* site used for pool construction (Fig. 1) is underlined. Complementarity between the random region and the (constant) 5'-end of the RNA is shaded (classes I and V). Both of these classes are 5'-kinases; these regions may serve to bind the 5'-end in the active sites of the ribozymes. Sites of 2'-thiophosphorylation are shown as white letters in black boxes. Clone Kin.47 is inactive, and contains a G to A mutation at the site of 2'-thiophosphorylation. The sequences of the constant primer binding regions (Fig. 1) are not shown except for the first three bases following the 5' primer binding site (AGA). The length of the original pool (not including primer binding sites) was 138 nucleotides. Point deletions may have occurred during the chemical synthesis of the DNA pool, and larger deletions may be due to annealing of primers to sites in the random regions during reverse-transcription or PCR. **b**, Top: structure of the ATP aptamer consensus⁴. Conserved bases in the loop are shown in capital letters. Positions that tend to be A, but which can vary, are shown in lower case. The bulged G is also conserved, but the stem regions (aside from being base paired) and the right hand loop are not. Bottom: possible secondary structures of the ATP aptamer domains from four of the major classes of ribozymes. The sequence of the most active clone is shown in each case. Positions in the loop regions that differ from the consensus sequence for the ATP aptamer are highlighted in grey. One of the stem regions from each of the classes II, VI and VII is missing, and so these structures are not shown. **c**, Kinetic parameters for the best member from each of classes I–IV.

b **ATP APTAMER CONSENSUS**



METHODS. Cycle 13 pool RNA was cloned using the pT7 Blue T-Vector kit (Novagen). Rates for each clone were determined (Fig. 2 legend) at six different ATP- γ S concentrations, ranging from 20 μM –2.5 mM. The data were fitted to the Michaelis–Menten equation³⁹ using the program Kaleidagraph (Abelbeck Software). Multiple values are independent experiments. The errors are standard errors of fit to the Michaelis–Menten equation.

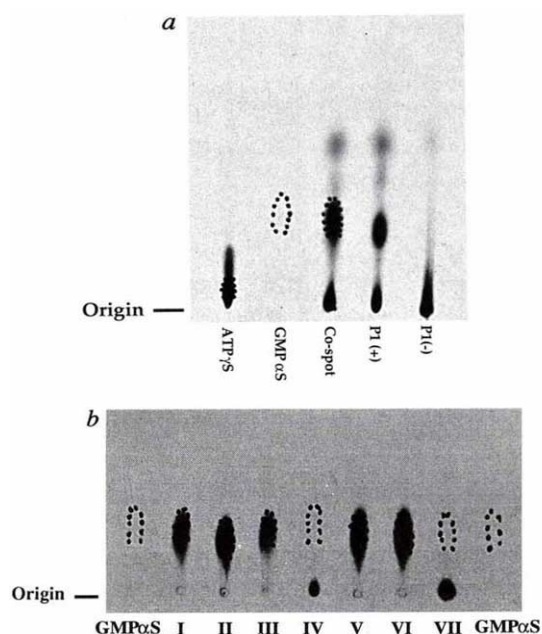


FIG. 4 Two different reactions are catalysed by selected RNA. *a*, Analysis of pool 13 RNA reaction products. Nuclease P1 digests of pool 13 RNA that had been allowed to react with ATP- γ - ^{35}S were run on polyethyl- enimine (PEI) cellulose TLC plates in a solvent system that resolves all four nucleotide 5'-monophosphates. Nuclease P1 digests RNA to nucleotide 5'-monophosphates, and thus if the 5'-guanosine is thiophosphorylated phosphorylated by the ribozyme, this treatment should yield ^{35}S -labelled GMP- α S. From left to right are shown: ATP- γ - ^{35}S marker; GMP- α S marker; GMP- α S plus the nuclease P1 digest of autothiophosphorylated pool 13 RNA; nuclease P1 digest of autothiophosphorylated pool 13 RNA; undigested autothiophosphorylated pool 13 RNA. Quantitation shows that 50% of pool 13 yields a product that comigrates with GMP- α S, indicating that approximately this fraction of the active molecules in the pool are 5'-kinases. *b*, PEI cellulose TLC of nuclease P1 digests of the most active clones from each of the seven major classes of ribozymes, co-spotted with GMP- α S. All of the other members of each class showed the same behaviour as the most active clones.

METHODS. RNA (1 μM) was allowed to react at 23 °C with $\sim 1 \mu\text{M}$ ATP- γ - ^{35}S in reaction buffer (Fig. 1) for 4–18 h. The RNA was then separated from nucleotides by G-50 spin-column gel filtration chromatography (Boehringer). The RNA was digested with nuclease P1 as described elsewhere^{24,26}. An aliquot was then spotted directly on a PEI cellulose TLC plate and developed in 1 M LiCl, 10 mM DTT²⁴. The products were localized by UV shadowing (for unlabelled GMP- α S) or by autoradiography. Thiophosphate containing nucleotides run slower in this system than do the corresponding phospho-nucleotides, presumably because there is weaker interaction between Li^+ and the thiophosphate than there is with the phosphate.

form base pairs with the ribozyme's substrate binding site, leading to very low K_m values. A comparison between the kinase ribozymes described here and the self-cleavage reaction catalysed by the *Tetrahymena* group I intron is particularly relevant, however, because both RNAs use external small molecule substrates (ATP- γ S and guanosine nucleotides, respectively) to modify themselves. The best kinase that we have characterized (Kin.25; class II) phosphorylates itself with a k_{cat} of approximately 0.3 min^{-1} and a k_{cat}/K_m of $6 \times 10^3 \text{ min}^{-1} \text{ M}^{-1}$. The *Tetrahymena* self-splicing intron has a k_{cat} of 0.5 min^{-1} and a k_{cat}/K_m of $2.5 \times 10^4 \text{ min}^{-1} \text{ M}^{-1}$ (ref. 28). Thus, from a vanishingly small sampling of sequence space, we have been able to isolate a molecule with catalytic activity essentially as good as that of a ribozyme found in nature (although, it should be noted that the reaction catalysed by the *Tetrahymena* intron is chemically more

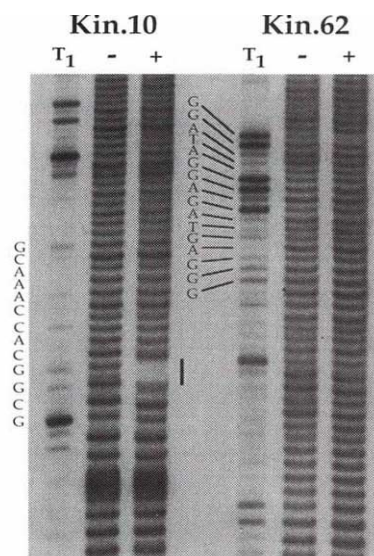


FIG. 5 Partial alkaline hydrolysis of thiophosphorylated class IV and VII RNAs. (–) Lanes: RNA was ^{32}P end-labelled with polynucleotide kinase and subjected to alkaline hydrolysis. (+) Lanes: RNA was allowed to autothiophosphorylate, affinity-purified on thiopyridine-activated thiopropyl Sepharose, 5'-end labelled with ^{32}P , and then base-hydrolysed. Gaps (highlighted by vertical black bars) are seen at G62 and G83 for Kin.10 and Kin.62 RNA, respectively, corresponding to 2'-thiophosphorylation at these positions. The faint band in the gap at G83 in the Kin.62 ladder may be due to hydrolysis of the thiophosphate during the base treatment. RNase T1 ladders of 5'-end labelled RNAs from each clone were run as sequence markers.

METHODS. RNA was reacted with ATP- γ S as described in the legend to Fig. 4, except that 100 μM unlabelled ATP- γ S was used. The thiophosphorylated RNAs were purified on thiopyridine-activated thiopropyl Sepharose, and then 5'-end labelled using T4 polynucleotide kinase and ATP- γ - ^{32}P . Alkaline hydrolysis was performed in 50 mM sodium carbonate/bicarbonate buffer, pH 9.0, 0.1 mM EDTA for 3 min at 90 °C. Reaction products were analysed on an 8% polyacrylamide/8 M urea gel.

difficult than that catalysed by the kinases). The activities of our ribozymes have not yet begun to approach those of the corresponding protein enzymes, however; the k_{cat} for T4 polynucleotide kinase is $25,000 \text{ min}^{-1}$ and the k_{cat}/K_m is approximately $6 \times 10^8 \text{ min}^{-1} \text{ M}^{-1}$ (K_m for ATP)³³.

Class I–IV kinases show specificity for ATP- γ S as a substrate. No reaction ($<0.1\%$ ATP- γ S rate) could be detected with GTP- γ S, indicating that the RNAs can discriminate between similar substrates. Interestingly, as much as 30% of the cycle 13 pool RNA can use GTP- γ S as a substrate, and thus pool 13 does contain molecules with less stringent substrate specificities. The class I–IV kinases are also able to discriminate between ATP- γ S and ATP ($k_{\text{obs}}(\text{ATP-}\gamma\text{S})/k_{\text{obs}}(\text{ATP})$: class I, 55; class II, 300; class III, 150; class IV, ≥ 300 ; 100 μM ATP, ATP- γ S). As these values are significantly larger than the three to ten fold higher intrinsic reactivity of ATP- γ S relative to ATP^{34,35}, the data suggest that the thiophosphate is important for binding, catalysis or both. Furthermore, pool 13 RNA is not detectably labelled by either [α - ^{35}S]ATP or [α - ^{32}P]ATP suggesting that 5'-splicing is not a reaction that occurs in the pool (unless the γ -thiophosphate is an absolute requirement for the molecules that carry out this reaction).

Rate acceleration

We have been unable to detect the uncatalysed background reaction for the thiophosphorylation of RNA (or guanosine) by ATP- γ S. Based on the sensitivity of these experiments, we esti-

mate the lower limit for the rate acceleration of the kinase ribozymes to be 10^5 -fold. At 70°C the rate of hydrolysis of ATP in the presence of Mg^{2+} is $\sim 4 \times 10^{-4} \text{ min}^{-1}$ (pH 6–8)³⁶. Correcting for the temperature (in general, increasing temperature 10°C produces $\sim 2\times$ increase in rate) and 55 M water, this value gives a second order rate constant of $\sim 5 \times 10^{-7} \text{ min}^{-1} \text{ M}^{-1}$. Assuming triphosphates behave like phosphate monoesters, ATP- γS should hydrolyse 3–10 times faster than ATP^{34,35}. Taking this factor into account, we estimate the rate enhancement of our ribozymes, $[k_{\text{cat}}/K_m]/[k_{\text{hydrolysis}}]$, would be $6 \times 10^3 \text{ min}^{-1} \text{ M}^{-1}/\sim 5 \times 10^{-6} \text{ min}^{-1} \text{ M}^{-1}$ or 10^9 -fold. This enhancement corresponds to an effective molarity of 10^4 – 10^5 M for ATP in the ATP-ribozyme complex ($k_{\text{cat}}/k_{\text{hydrolysis}} = 0.3 \text{ min}^{-1}/5 \times 10^{-6} \text{ min}^{-1} \text{ M}^{-1}$).

Intermolecular catalysis and turnover

At least one of our selected kinases is capable of catalysing the thiophosphorylation of a separate RNA substrate. Kin.46 (class I) can transfer the γ -thiophosphate from ATP- γS to the 5'-end of short oligoribonucleotides with the same sequence as the 5'-end of the ribozyme (Fig. 6). Full-length Kin.46 catalyses this reaction ~ 500 -fold more slowly than the autothiophosphorylation reaction. Part of the reason for the decreased activity is likely to be competition for the active site between the 5'-end of the RNA and the exogenous oligoribonucleotide substrate. When the 5'-constant region of the RNA is removed (via PCR with an internal 5'-primer, followed by transcription), the activity increases >100 -fold to a level only slightly lower than that of the autothiophosphorylation reaction. (At saturating concentrations of the 6-mer 5'-GGAACC-3' (100 μM) and ATP- γS (2.5 mM) the initial rate of thiophosphorylation is $0.05 \mu\text{M min}^{-1}$ with 1 μM ribozyme. Under the same conditions, the 7-mer 5'-GGAACCU-3' is thiophosphorylated with an initial rate of $0.17 \mu\text{M min}^{-1}$. In comparison, the rate of autothiophosphorylation for full length Kin.46 RNA (1 μM) with 2.5 mM ATP- γS is $0.3 \mu\text{M min}^{-1}$.) At 23°C the ribozyme performs 10 turnovers per hour with the 7-mer substrate, and is thus acting as a true enzyme.

Discussion

We have shown that RNA is capable of catalysing, with surprising efficiency, reactions directly analogous to those performed by protein polynucleotide kinases. In fact, in a vanishingly small sampling of sequence space, we have found at least five different RNAs capable of acting as 5'-kinases. At least two different 2'-kinases were also found in this small fraction of sequence space.

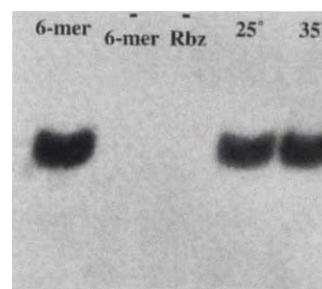


FIG. 6 Class I RNAs act as true enzymes. Transfer of thiophosphate from ATP- γS to a 6-mer oligoribonucleotide by a 5'-truncated form of Kin.46 RNA. The 6-mer had the same sequence as the 5'-end of the full-length Kin.46 RNA. 6-mer lanes: 5'-³⁵S-thiophosphate-labelled 6-mer run as a marker. -Rbz: no ribozyme control. -6-mer: no 6-mer RNA control. 25°/35°: 24-h time points of reactions done at the indicated temperatures. 35 $^\circ\text{C}$ was found to be the optimal temperature for the reaction.

METHODS. 1 μM ribozyme was incubated as described in the legend to Fig. 2 except that 2.5 mM ATP- γS was used, and 100 μM 5'-HO-GGAACC-3' RNA was added. The ribozyme had the same sequence as Kin.46, except that the 5'-constant region had been deleted. The 6-mer was synthesized by run-off transcription²³ and was dephosphorylated with calf intestinal alkaline phosphatase before ion-exchange HPLC purification. The thiophosphorylated 6-mer marker was made by end-labelling 5'-GGAACC-3' with ATP- γ -³⁵S using T4 polynucleotide kinase. Products were analysed on 20% polyacrylamide/8 M urea gels.

These results imply that in sequence space as a whole there may be many more active molecules, some with far greater catalytic power than those isolated in this study. We suspect, however, that by beginning our selection with a pool of molecules already containing a substrate binding site, we greatly increased the odds of finding catalytically active sequences. This technique should be generally applicable to the selection of novel ribozymes that use other small molecule substrates and cofactors.

Similar evolutionary mechanisms may have occurred in nature. For example, the evolution of one function, such as a polynucleotide kinase, could have facilitated the subsequent evolution of related functions, such as an ATP-dependent ligase (or any other ribozyme that used ATP as a substrate) by the mixing, matching and modification of small functional domains. This is precisely the mechanism proposed by Gilbert³⁷ for the evolution of new catalytic functions in the RNA, and later the protein, worlds. □

Received 12 May; accepted 12 July 1994.

1. Ellington, A. D. & Szostak, J. W. *Nature* **346**, 818–822 (1990).
2. Ellington, A. D. & Szostak, J. W. *Nature* **355**, 850–852 (1992).
3. Famulok, M. & Szostak, J. W. *J. Am. chem. Soc.* **114**, 3990–3991 (1992).
4. Sassanfar, M. & Szostak, J. W. *Nature* **364**, 550–552 (1993).
5. Connel, G. J., Illangsekare, M. & Yarus, M. *Biochemistry* **32**, 5497–5502 (1993).
6. Lorsch, J. R. & Szostak, J. W. *Biochemistry* **33**, 973–982 (1994).
7. Famulok, M. *J. Am. chem. Soc.* **116**, 1698–1706 (1994).
8. Jenison, R. D., Gill, S. C., Pardi, A. & Polisky, B. *Science* **263**, 1425–1429 (1994).
9. Bartel, D., Zapp, M. L., Green, M. R. & Szostak, J. W. *Cell* **67**, 529–536 (1991).
10. Tuerk, C. & Gold, L. *Science* **249**, 505–510 (1990).
11. Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H. & Toole, J. J. *Nature* **355**, 564–566 (1992).
12. Tsai, D. E., Kenan, D. J. & Keene, J. D. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8864–8868 (1992).
13. Beaudry, A. A. & Joyce, G. F. *Science* **257**, 635–641 (1992).
14. Green, R. & Szostak, J. W. *Science* **258**, 1910–1915 (1992).
15. Lehman, N. & Joyce, G. F. *Nature* **361**, 182–185 (1993).
16. Pan, T. & Uhlenbeck, O. C. *Biochemistry* **31**, 3887–3895 (1992).
17. Bartel, D. P. & Szostak, J. W. *Science* **261**, 1411–1418 (1993).
18. Prudent, J. R., Uno, T. & Schultz, P. G. *Science* **264**, 1924–1927 (1994).
19. Richardson, C. C. in *The Enzymes* Vol. 14 (ed. Boyer, P.) 299–314 (Academic, New York, 1981).
20. Zimmerman, S. B. & Pfeiffer, B. H. in *The Enzymes* Vol. 14 (ed. Boyer, P.) 315–329 (Academic, New York, 1981).
21. Zaug, A. J. & Cech, T. R. *Biochemistry* **25**, 4478–4482 (1986).

22. Cadwell, R. C. & Joyce, G. F. *PCR Meth. Appl.* **2**, 28–33 (1992).
23. Milligan, J. F. & Uhlenbeck, O. C. *Meth. Enzym.* **180**, 51–64 (1989).
24. Westaway, S. K., Belford, H. G., Apostol, B. L., Abelson, J. & Greer, C. L. *J. biol. Chem.* **268**, 2435–2443 (1993).
25. Domdey, H. & Gross, H. J. *Analyt. Biochem.* **98**, 346–352 (1979).
26. Konarska, M., Filipowicz, W., Domdey, H. & Gross, H. J. *Nature* **293**, 112–116 (1981).
27. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. *Cell* **35**, 849–857 (1983).
28. Bass, B. & Cech, T. R. *Nature* **308**, 820–826 (1984).
29. Chowrira, B. M., Berzal-Herranz, A. & Burke, J. M. *Nature* **354**, 320–322 (1991).
30. Perrotta, A. T. & Been, M. D. *Biochemistry* **31**, 16–21 (1992).
31. Dahm, S. C., Derrick, W. B. & Uhlenbeck, O. C. *Biochemistry* **32**, 13040–13045 (1993).
32. Pyle, A. M. & Green, J. B. *Biochemistry* **33**, 2716–2725 (1994).
33. Lillehaug, J. R. & Kleppe, K. *Biochemistry* **14**, 1221–1225 (1975).
34. Herschlag, D., Piccirilli, J. A. & Cech, T. R. *Biochemistry* **30**, 4844–4854 (1991).
35. Breslow, R. & Katz, I. J. *J. Am. chem. Soc.* **90**, 7376–7377 (1968).
36. Ramirez, F., Marecek, J. F. & Szamosi, J. J. *J. Org. Chem.* **45**, 4748–4752 (1980).
37. Gilbert, W. *Nature* **319**, 618 (1986).
38. Eckstein, F. A. *Rev. Biochem.* **54**, 367–402 (1985).
39. Fersht, A. *Enzyme Structure and Mechanism* (Freeman, New York, 1985).

ACKNOWLEDGEMENTS. We thank members of our laboratory for critically reading the manuscript; D. Bartel, A. Hager, C. Lauhon, R. Rohatgi and C. Wilson for advice and discussion; and D. Herschlag for advice on calculating rate enhancements and for comments on the manuscript. J.R.L. was supported by an NIH training grant to the Departments of Biochemistry and Molecular Biology and Cellular and Developmental Biology at Harvard University. This work was supported by grants from the NIH and Hoechst AG.

Structure of influenza haemagglutinin at the pH of membrane fusion

Per A. Bullough^{*†}, Frederick M. Hughson^{*}, John J. Skehel[‡] & Don C. Wiley^{*§}

^{*} Department of Biochemistry and Molecular Biology and [§] Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

[‡] National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Low pH induces a conformational change in the influenza virus haemagglutinin, which then mediates fusion of the viral and host cell membranes. The three-dimensional structure of a fragment of the haemagglutinin in this conformation reveals a major refolding of the secondary and tertiary structure of the molecule. The apolar fusion peptide moves at least 100 Å to one tip of the molecule. At the other end a helical segment unfolds, a subdomain relocates reversing the chain direction, and part of the structure becomes disordered.

INFECTION by enveloped viruses involves fusion of viral and cellular membranes with subsequent transfer of viral genetic material into the cell. The virus components that mediate fusion are membrane glycoproteins, among the best characterized of which is the influenza virus haemagglutinin (HA)¹. The HA is also responsible for binding influenza viruses to their sialylated cell-surface receptors, following which bound virus is internalized by endocytosis. At the low pH of endosomes, between pH 5 and pH 6, the fusion potential of the HA is activated^{2–4} in a process requiring structural changes in HA^{5,6} (reviewed in ref. 1). We report here the results of crystallographic analyses of a soluble fragment from low-pH-treated HA which indicate that the fusion-pH-induced conformation is substantially different from the neutral pH conformation.

Native HA has a relative molecular mass of 220K and is a trimer of identical subunits, each of which comprises two glycopolypeptides linked by a single disulphide bond, HA₁ (328 residues) and HA₂ (221 residues). Each HA₂ chain is anchored at its carboxy terminus in the viral membrane. Soluble trimers can be released by treating virus with the protease bromelain, which cleaves each HA₂ chain once after residue 175 (refs 7, 8). The structure^{9,10} of this trimer, BHA (Fig. 1a), reveals that the HA₂ chains are major components of a mainly α -helical stem domain which forms the centre of the molecule. The HA₁ chains also contribute to the stem structure but primarily form three membrane distal globular domains containing the receptor binding sites. This structure poses two difficulties in understanding the role of the HA in membrane fusion⁹. First, the "135 Å" length of the molecule seems to restrict membranes from any closer approach. Second, the conserved, hydrophobic sequence at the N terminus of HA₂, the so called 'fusion peptide'^{11,12}, is ~100 Å from the distal tip and ~35 Å from the viral membrane end of the molecule (Fig. 1b); a major conformational change would be required for this segment to approach either the virus or host cell membrane.

When incubated at the pH of fusion, BHA undergoes structural changes, two consequences of which are particularly important for investigations of the fusion-activated molecule. First, it aggregates through exposure of the fusion peptide^{11,13,14}. Such aggregates are unsuitable for crystallographic studies. Second, it becomes highly susceptible to proteases. In particular, thermolysin cleaves near the fusion peptide and the removal of this apolar segment solubilizes the fusion-pH-induced protein aggregate¹⁵. The soluble trimeric fragment considered here,

called TBHA₂, is prepared from BHA at pH 5.0 by successive digestion with trypsin and thermolysin^{13,15,16}. Each monomer includes residues 38–175 of the HA₂ chain disulphide linked to residues 1–27 of the HA₁ chain.

Dramatic differences are observed between the BHA and TBHA₂ structures. Large changes occur at both ends of the HA₂ chain. The N terminus is displaced 100 Å, a movement which in the intact molecule could transport the fusion peptide 150 Å or more. This movement appears to be driven by recruitment of at least 36 additional residues to the viral membrane-distal end of the triple-stranded α -helical coiled coil of BHA. At the C-terminal end, seven residues in the middle of the long BHA α -helix refold to form a bend, allowing the remainder of the helix and three short β -strands packed against it to jackknife back and pack against the start of the helix, a reorientation of 180°. These two structural changes displace most of the residues against which the C-terminal region of the HA₂ chain, residues 140–175, was packed in native BHA, apparently extruding it into an extended and partially disordered structure.

Structure determination

The TBHA₂ structure was determined using crystals cooled to –170 °C in cryoprotectant buffer¹⁶ (Table 1). Briefly, a multiple isomorphous replacement (MIR) map calculated to 5.0 Å resolution was subjected to iterative cycles of threefold non-crystallographic symmetry averaging and phase extension; single isomorphous replacement (SIR) phases from the platinum derivative were combined with extended phases at each resolution step to 3.5 Å. An initial model built into this MIR/SIR map was improved by further rounds of building, averaging, and refinement at progressively higher resolution, guided chiefly by improvement in the free *R* factor¹⁷. Details are given in Table 1.

The current model (Fig. 2) includes residues 12₁–16₁ and 40₂–153₂ of the first monomer, residues 11₁–16₁ and 40₂–162₂ of the second monomer, residues 10₁–17₁ and 40₂–162₂ of the third monomer, and 37 water molecules (numbering scheme is that for intact HA: HA₁ and HA₂ residues are distinguished by the appropriate subscript). The chain trace is clear throughout this modelled region, as is electron density for most (89%) of the side chains, as judged by simulated annealing omit maps. However, a significant fraction (23%) of the molecule, including from two to 22 residues at each of the twelve chain termini, has not been modelled for lack of connected electron density. These apparently disordered or partially ordered regions pose difficulties for further refinement. In at least one case (residues 154₂–162₂), residues that are well-ordered in two monomers are

[†] Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

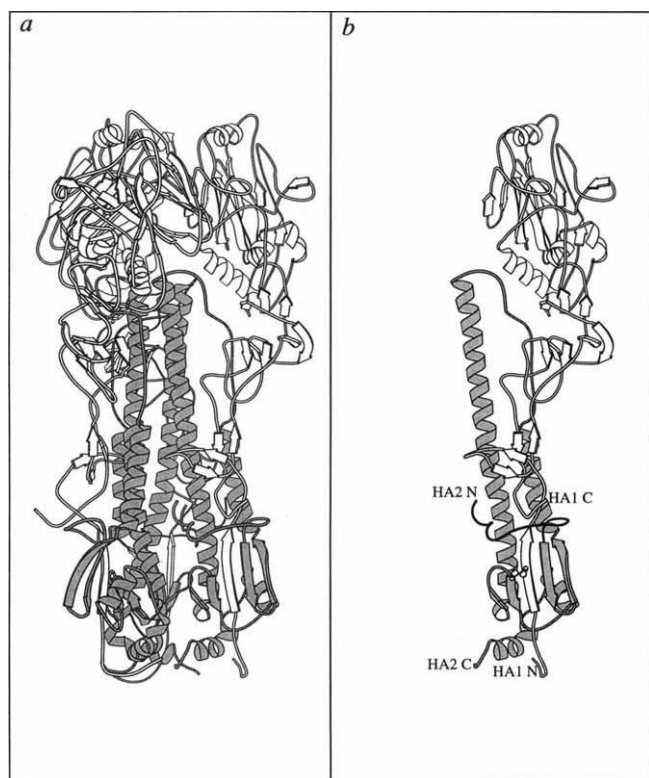


FIG. 1 BHA structure¹⁰. *a*, The BHA trimer and *b*, monomer are shown, with the HA₁ chains white and the HA₂ chains shaded. In *b*, the locations of the fusion peptide (black) and the N and C termini of the chains are indicated. Also shown is the interchain disulphide bond. The viral membrane end of the molecule is at the base, near the HA₂ C terminus generated by proteolytic removal of the transmembrane anchor sequence; the viral membrane distal globular domains are at the top. Figs 1, 2, 3c and *d* were generated using MOLSCRIPT⁴⁹

prevented from adopting this conformation in the third monomer by a crystal contact, and indeed the electron density for these residues is uninterpretable.

The major feature of the TBHA₂ structure (Fig. 2) is the three-stranded α -helical coiled coil, approximately 100 Å long, comprising one long α -helix from each monomer (residues 40₂–105₂). This is followed in each monomer by a small connecting loop and a shorter helix (113₂–129₂) which packs in an antiparallel orientation against the long helix of the same monomer and that of another monomer. A β -hairpin (131₂–140₂), together with an additional strand contributed by HA₁ residues (11₁–16₁), forms a small antiparallel β -sheet. HA₁ is connected to HA₂ by a disulphide bond between residues 14₁ and 137₂. An additional short helix (146₂–153₂) packs against the coiled coil. Beyond residue 153₂ for one chain, and beyond residue 162₂ for the other two chains, the electron density is not interpretable, presumably indicating disorder.

BHA and TBHA₂ compared

Only thirty residues (76₂–105₂; C in Fig. 3*a, b*) have the same structure in BHA and TBHA₂ (root mean square difference, 0.6 Å). In both molecules, these residues form part of the central triple-stranded α -helical coiled coil. For the purpose of comparing the two structures, BHA and TBHA₂ were overlaid by aligning these residues.

Secondary structure changes are summarized in Fig. 3*a*. Four α -helical segments (A, C, D, G) and three β -strands (1, E, F) are conserved in both structures (Fig. 3). An irregular segment of BHA, the extended loop B between helices A and C, adopts a helical conformation in TBHA₂. Conversely, a segment in BHA in the middle of the long α -helix (106₂–112₂) refolds into

TABLE 1 Crystallographic and refinement statistics

Data collection and MIR statistics			
Parameter	Native	TAMM*	K ₂ PtCl ₄
Resolution (Å)	15.0–2.5	15.0–5.0	15.0–3.2
Completeness ($I > \sigma(I)$)	0.86	0.78	0.92
$R_{\text{merge}}^{\dagger}$	0.065	0.083	0.088
$R_{\text{iso}}^{\ddagger}$	—	0.28	0.15
Number of sites	—	2	3
$R_{\text{c}}^{\S}$	—	0.59	0.60
Phasing power $\parallel$	—	0.97	1.16
Refinement statistics (6.0–2.50 Å)			
Number of atoms*	3,135		
Number of reflections (working set)	25,888		
$R_{\text{cryst}}^{\#}$	0.222		
Number of reflections (free set)	1,905		
$R_{\text{free}}^{\#}$	0.291		
Deviations from ideal geometry	0.015 Å (bonds)		
	2.17° (angles)		
Real space fit**	0.88 ± 0.09 (main chain)		
	0.82 ± 0.12 (side chain)		

Crystals in space group C222₁ ($a = 168.7$ Å $b = 231.7$ Å $c = 53.8$ Å) were grown as described¹⁶. Native X-ray data were collected on phosphor image plates at the Cornell High Energy Synchrotron Source (CHESS) from a single crystal in buffer plus 20% glycerol¹⁶ and integrated using DENZO (Z. Otwinowski, personal communication). Derivative data were collected from two crystals soaked in buffer plus glycerol and 0.1 mM TAMM or 1.0 mM K₂PtCl₄ for one day. Data were recorded using an Elliot GX-13 X-ray source and either a Xenotronics area detector (TAMM derivative) or a Mar Research Image Plate (K₂PtCl₄ derivative). Data were integrated with BUDDHA³⁸ (TAMM derivative) or MOSFLM³⁹ (K₂PtCl₄ derivative). All data were reduced and scaled using CCP4 programs⁴⁰. One Pt site was located in a difference Patterson map; two additional Pt sites and two Hg sites were observed in difference Fourier maps and confirmed in difference Patterson maps. Heavy-atom sites were refined using HEAVY⁴¹. The three Pt sites defined the approximate position and orientation of one non-crystallographic molecular three-fold symmetry axis in each asymmetric unit, which was refined using CYLINDER⁴². Iterative averaging/solvent flattening⁴³ produced a map which was partially fitted with largely α -helical segments of polyaniline using the programme O⁴⁴. Heavy atom positions were re-refined against model phases to generate an improved MIR map, calculated at 5 Å, which was extended to 3.5 Å by means of real-space averaging, phase extension, and phase combination (see text). Cycles of rebuilding, positional refinement with strict non-crystallographic symmetry (using XPLOR⁴⁵), and averaging (using RAVE and associated programmes (G. J. Kleywegt and T. A. Jones, personal communication)) at progressively higher resolution were used to improve the model as judged by the free R factor¹⁷. Individual isotropic B factors were introduced and refined. When no further improvement could be obtained, the model was rebuilt into unaveraged maps. Cycles of building and refinement of the model without non-crystallographic symmetry constraints were performed (r.m.s. differences between monomer pairs (all atoms) are currently 1.9, 1.7 and 1.7 Å). Thirty-seven water molecules per trimer were added (4σ difference density, $B < 80$ Å², good geometry). We have not been able to extend the model into the disconnected density visible near some chain termini. Despite high B factors (Fig. 2 legend), the chain trace is clear throughout in simulated annealing omit maps, although one slight break in main chain density is seen (in the second monomer at the Phe 138₂ nitrogen) and density for five side chains is absent. Neither three-dimensional profile analysis⁴⁶ nor analysis of nonbonded atomic interactions⁴⁷ identifies any significant problem with the current model, and less than 3% of the (ϕ , ψ) angles lie outside allowed regions in a Ramachandran plot. Refined coordinates will be deposited in the Protein Data Bank (Brookhaven National Laboratory, Upton, NY).

* TAMM, tetrakis(acetoxymethyl)mercuric methanolate.

$\dagger R_{\text{merge}} = \sum_i |I_i - \langle I \rangle| / \sum_i \langle I \rangle$ where I_i is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

$\ddagger R_{\text{iso}} = \sum_i |F_{\text{ph}}| - |F_{\text{p}}| / \sum_i |F_{\text{p}}|$ where F_{ph} and F_{p} are the derivative and native structure factors respectively.

$\S R_{\text{c}} = \sum_i ||F_{\text{ph}} \pm F_{\text{p}}| - |F_{\text{calc}}|| / \sum_i |F_{\text{ph}} \pm F_{\text{p}}|$ for centric reflections where F_{calc} is the calculated heavy-atom structure factor.

$\parallel$ Phasing power = $\langle F_{\text{h}} \rangle / E$ where $\langle F_{\text{h}} \rangle$ is the root-mean-square heavy-atom structure factor and E is the residual lack of closure error.

$\#$ Non-hydrogen atoms included in the current model.

$\# R_{\text{cryst, free}} = \sum_i |F_{\text{obs}}| - |F_{\text{calc}}| / \sum_i |F_{\text{obs}}|$, where the crystallographic and free R factors are calculated using the working and free reflection sets respectively ($F_{\text{obs}} > 2\sigma(F_{\text{obs}})$). The free reflections (about 7% of the total) were chosen before averaging and phase extension and were held aside throughout refinement.

** Real-space fit⁴⁴ of the current model to a SIGMAA-weighted $2F_{\text{obs}} - F_{\text{calc}}$ map⁴⁸ using phases calculated from the current model.

a loop, allowing a 180° bend between helices C and D in TBHA₂. An α -helix, H, in BHA (159₂–170₂) and five residues beyond it, 171₂–175₂, appear to be disordered in TBHA₂. The overall helix content agrees well with that measured by circular dichroism¹⁸ and, because of compensating changes, is very similar to that of the corresponding portions of the BHA structure.

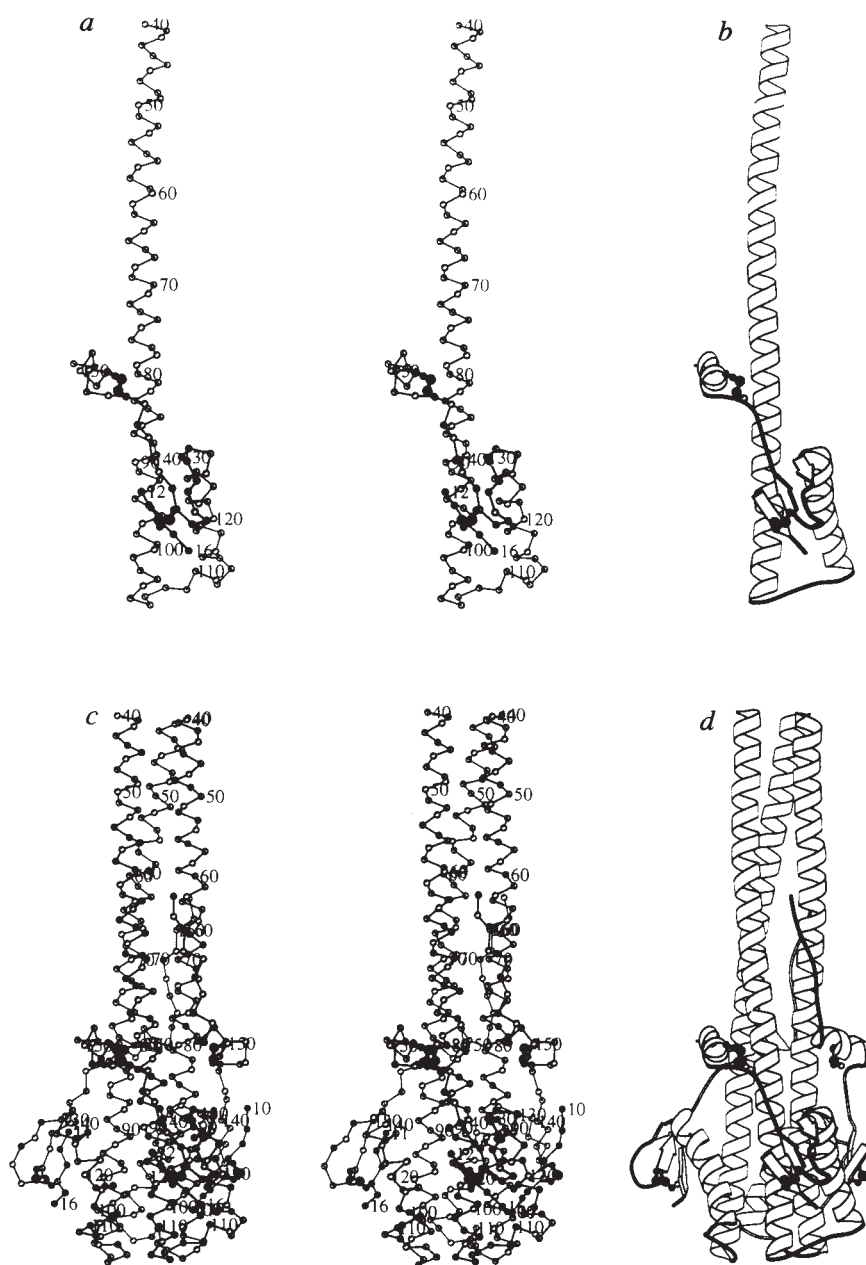


FIG. 2 TBHA₂ structure. *a*, Stereo image and *b*, ribbon diagram of the monomer, showing residues 12–16 of HA₁ and residues 40–153 of HA₂. *c*, Stereo image and *d*, ribbon drawing of the trimer. The first monomer is as in *a* and *b*; the second contains HA₁ residues 11–16 and HA₂ residues 40–162, and the third contains HA₁ residues 10–17 and HA₂ residues 40–162. In *a* and *c*, HA₁ is shown with filled balls and HA₂ with open balls. Side chains are shown for cysteine residues, all of which participate in disulphide bonds. The mean *B* factor for all model atoms is 60 Å², in agreement with the *B* factor calculated from a Wilson plot (58 Å²). Two short regions, 130₂–136₂ and 150₂–160₂, show particular evidence of disorder, with mean *B* > 75 Å² for each residue.

Tertiary structure refolding

N-terminal. Major refolding of secondary structure elements occurs at both ends of the TBHA₂ molecule. Helix A (38₂–55₂), which packs against the central helix CD in BHA, becomes the top of the triple-stranded α -helical coiled coil at the central core of TBHA₂ (Fig. 3*b–d*). Although the N-terminal residues (1₂–37₂) have been removed in the preparation of TBHA₂, the location of helix A suggests that the fusion peptide (1₂–23₂) and two β -strands (24₂–37₂) would also relocate to this end of the intact molecule, a distance of 100 Å or more. Loop B of BHA, which was docked against the C region of the long helix and made extensive contacts with HA₁, adopts an α -helical conformation and joins the triple-stranded coiled coil in TBHA₂ (Fig. 3*b–d*). This transition was anticipated¹⁹ by the observation that residues 38₂–125₂ contain heptad repeats of hydrophobic residues which could potentially form a coiled coil^{19,20}. It has been shown¹⁹ that peptides corresponding to B and the first four turns of C, with or without A, form α -helical trimers, demonstrating the energetic preference of this segment to be a coiled coil.

C-terminal. Nearer the C-terminal end of the molecule, a portion of helix CD (106₂–112₂) that made a number of contacts with the fusion peptide in BHA^{9,10} refolds to form a loop (Fig. 3*b–d*). In BHA, the bottom half (D) of the long α -helix diverges from the central axis of the molecule to form a tripod-like structure apparently stabilized by polar and nonpolar interactions with the fusion peptide. Below residue 112₂, the residues on the inner face of the long helices which surround the trimer axis are exclusively polar and/or charged. The relocation of the fusion peptide from the middle region of the CD helix may destabilize this tripod arrangement. Hence, although helix D (113₂–126₂) remains intact, it folds back together with the small antiparallel β -sheet (1, E, F in Fig. 3*b*) to pack against the triple-stranded coiled coil in an inverted orientation (Fig. 3*b–d*). There are many small changes in the architecture of this subdomain (1, D, E, F in Fig. 3*b*; blue in Figs 3*c, d*); its root-mean-square *Ca* coordinate difference versus BHA is over 3 Å. An extensive new hydrophobic core, which presumably stabilizes the fusion-pH-induced conformation, is formed as a result of this subdomain movement (Fig. 4*a*). The single Trp in each monomer (residue 92₂) is

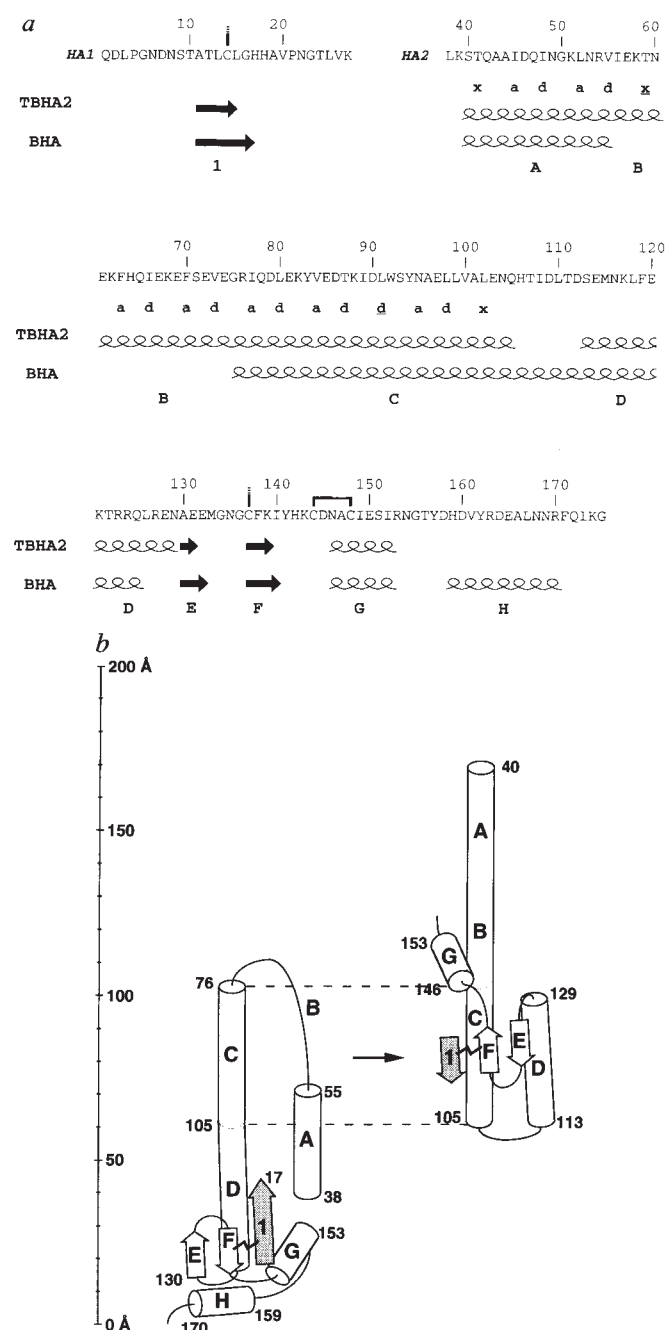
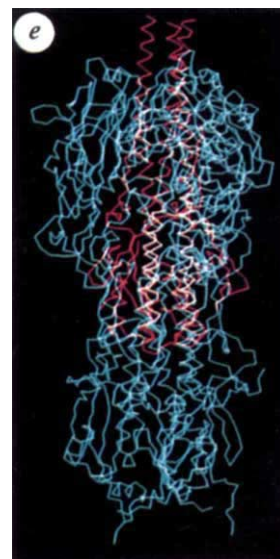
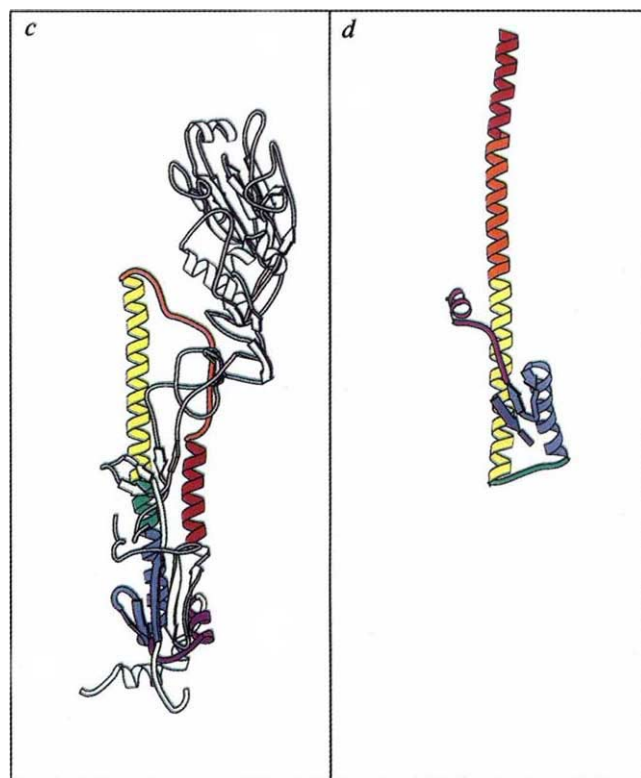


FIG. 3 The low-pH-induced conformational change. *a*, Secondary structure for TBHA₂ and the corresponding portion of BHA. α -helices and β -strands (arrows) are indicated below the amino-acid sequence. Letters A–H correspond to those in *b*. Core residue positions (a, d or x) for the TBHA₂ triple-stranded α -helical coiled coil are also shown (see text and Fig. 5); shifts in heptad register are centred on the underlined positions, as explained in the legend to Fig. 5. *b*, The structure of a TBHA₂ monomer (right) in schematic form is compared to the corresponding region of a BHA monomer (left). Consecutive regions of the HA₂ chain are labelled A–H (see text); H is apparently disordered in TBHA₂. The first β -strand of HA₁ is also shown, labelled '1'; both ends of this strand are apparently disordered in TBHA₂. The disulphide bond linking 141 and 137₂ is indicated. The two structures are aligned on



the C region, which is unaffected by the conformational change. *c*, *d*, The structure of a TBHA₂ monomer (*d*) is compared to that of an intact BHA monomer (*c*). Structural elements of the HA₂ chain of TBHA₂ are coloured in rainbow order from amino to carboxy terminus; corresponding regions of BHA are coloured identically. In addition, the HA₁ β -strand which is disulphide-linked to HA₂ is blue; the blue elements (1, D, E, F in *b*) move approximately as a subdomain in the conformation change (see text). Regions of BHA which were proteolytically removed to generate TBHA₂ are shown in grey. Regions of BHA corresponding to apparently disordered regions of TBHA₂ are shown in white. *e*, TBHA₂ (red) overlaid on BHA (blue). The two structures are aligned using the C region of the coiled coil (*a*, *b*). Figs 3e and 5 were generated using the program O⁴⁴.

buried, consistent with fluorescence measurements of TBHA₂ in solution¹⁸.

Residues 141₂–175₂, including helices G and H, which in BHA form a compact unit adjacent to the five-stranded β -sheet and part of helix D, adopt in TBHA₂ a more extended, partially disordered conformation running antiparallel to the

coiled coil (Fig. 3*b*–*d*). This region (yellow in Fig. 4*b*) docks against residues (other colours in Fig. 4*b*) in BHA, many of which become buried, move to distant locations, or rearrange relative to one another in TBHA₂ (Fig. 4*b* legend). The destruction of the site which presumably had stabilized the 141₂–175₂ unit may account for the partial disordering of this

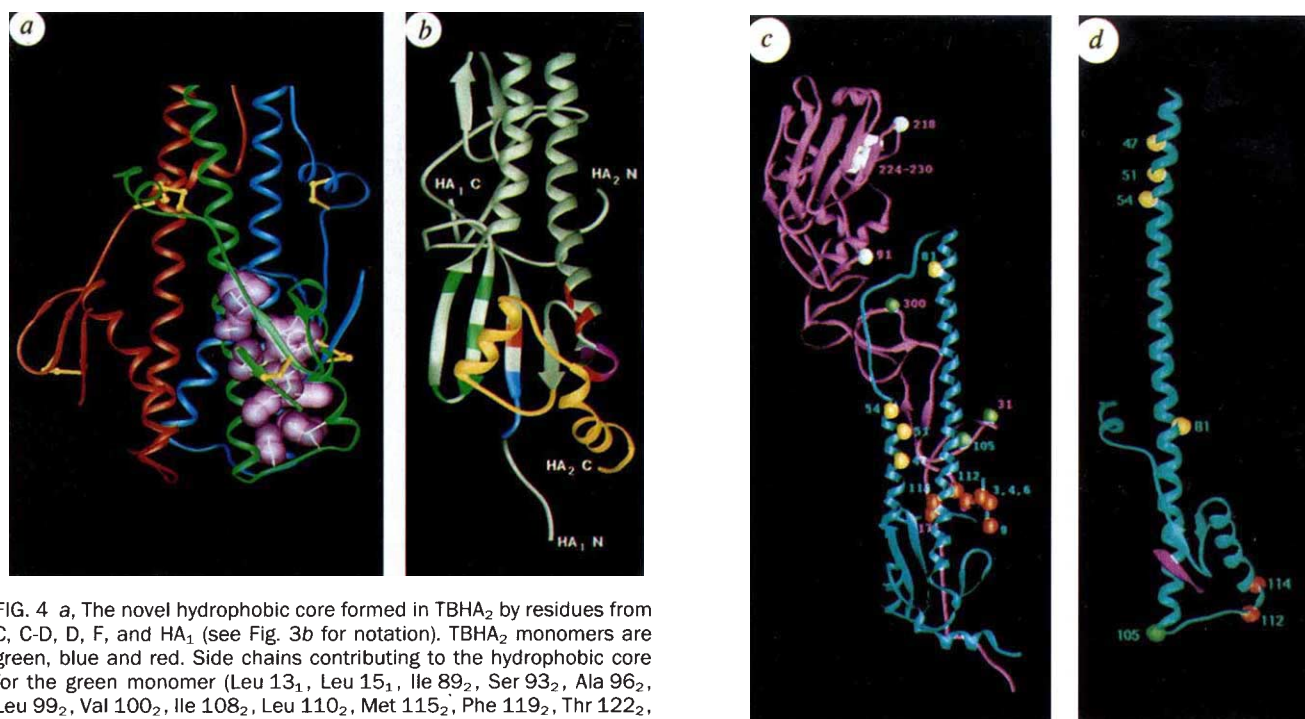


FIG. 4 a, The novel hydrophobic core formed in TBHA₂ by residues from C, C-D, D, F, and HA₁ (see Fig. 3b for notation). TBHA₂ monomers are green, blue and red. Side chains contributing to the hydrophobic core for the green monomer (Leu 13₁, Leu 15₁, Ile 89₂, Ser 93₂, Ala 96₂, Leu 99₂, Val 100₂, Ile 108₂, Leu 110₂, Met 115₂, Phe 119₂, Thr 122₂, Leu 126₂, Phe 138₂ and Ile 140₂) are shown in purple. b, The 'docking site' for the C-terminal portion of HA₂ in BHA. Residues 141₂–175₂ in BHA, including the G and H helices (Fig. 3b), are shown in yellow. Residues that make contacts with this region are green, red, magenta or blue; the remainder of the BHA monomer is grey. Green residues, in the N-terminal strands of HA₂, presumably move at least 100 Å as a result of the distal extension of the coiled coil at low pH. Red residues are largely buried in TBHA₂, inaccessible for interaction with the C-terminal region of HA₂. Magenta residues, although largely surface-exposed in TBHA₂, rearrange relative to one another. Blue residues are apparently disordered in at least one TBHA₂ monomer. c, d, Ribbon diagrams of BHA (c) and TBHA₂ (d) monomers showing the positions of

region in TBHA₂. Helix G (146₂–153₂) remains intact in TBHA₂, probably stabilized in part by the disulphide 144₂–148₂, but it docks against the C region of the long helix, approximately where the B loop had docked in BHA (Fig. 3b–d). Residues beyond 153₂ in one subunit and 162₂ in the others, including helix H, were not located in our electron density maps, presumably because they are disordered in TBHA₂.

HA₁–HA₂ interactions. In TBHA₂, HA₂ has refolded so extensively that many of the HA₁–HA₂ interactions seen in BHA would no longer be possible even if the bulk of the HA₁ chain were still present. In BHA, loop B (orange in Fig. 3c) makes all of the HA₁–HA₂ contacts involving the globular head domain of HA₁, and many involving the descending C-terminal HA₁ strand. Having refolded into a helical conformation in TBHA₂, the B region cannot participate in these interactions. Indeed, chemical crosslinking, antibody binding, proteolysis, electron microscopy, mutation, and site-specific modification data all indicate that the HA₁ globular domains in BHA or in the intact HA are displaced in the fusion pH conformational change^{5,6,13–15,21–25}. A second site (the C-D loop; green in Fig. 3c, d) which refolds significantly in TBHA₂ was the docking site for a loop (residues 24₁–37₁) in the ascending N-terminal HA₁ strand in BHA. Site-specific antipeptide antisera binding and proteolytic susceptibility at residue 27₁ indicate that the environment of this HA₁ loop rearranges in the intact protein at the pH of fusion^{5,11}.

The observed irreversibility of the fusion-pH-induced conformational transition^{11,26} and the increased thermostability of TBHA₂ compared to BHA even at neutral pH²⁷ suggest that the low-pH-induced structure is more thermodynamically stable

than the neutral pH structure, that is, that the neutral pH structure is metastable^{19,27}. Kinetically controlled activity and conformational refolding have also been observed in the serpin family of protease inhibitors (reviewed in ref. 28).

TBHA₂ coiled coil

The triple-stranded coiled coil of TBHA₂ displays, with few exceptions, the knobs-into-holes packing predicted by Crick²⁹ (Fig. 5a, b). The residues that compose the core are mostly hydrophobic and recur, in general, with a heptad (3–4) periodicity, where heptad positions are denoted a–g (Fig. 3a). At two positions along the coil an unusual 3–4–4–3 periodicity is observed (Figs 3a, 5c), as predicted in part by Ward and Doppeide²⁰ but not by others¹⁹, who assumed a single unbroken register. A small number of single-residue insertions or 'skips' are also observed in the heptad repeat sequences of proteins, such as myosin, which contain long double-stranded coiled coils³⁰. In TBHA₂, the net result of these shifts is to underwind the coiled coil to an average pitch of ~300–400 Å.

pH mutants

Mutant viruses were selected by Daniels *et al.* that undergo the conformational change and mediate fusion at elevated pH^{25,31}. The distribution of the mutations throughout the HA molecule, coupled with their effects on the pH of fusion, provide support both for the structural transition proposed here and for its functional significance in promoting membrane fusion. Specifically, these mutations can be seen to identify interactions which are modified in the transition state to the fusion-active structure. They have been placed into four groups (Fig. 4c, d). The first group (red in Fig. 4c, d) includes residues that appear to stabilize

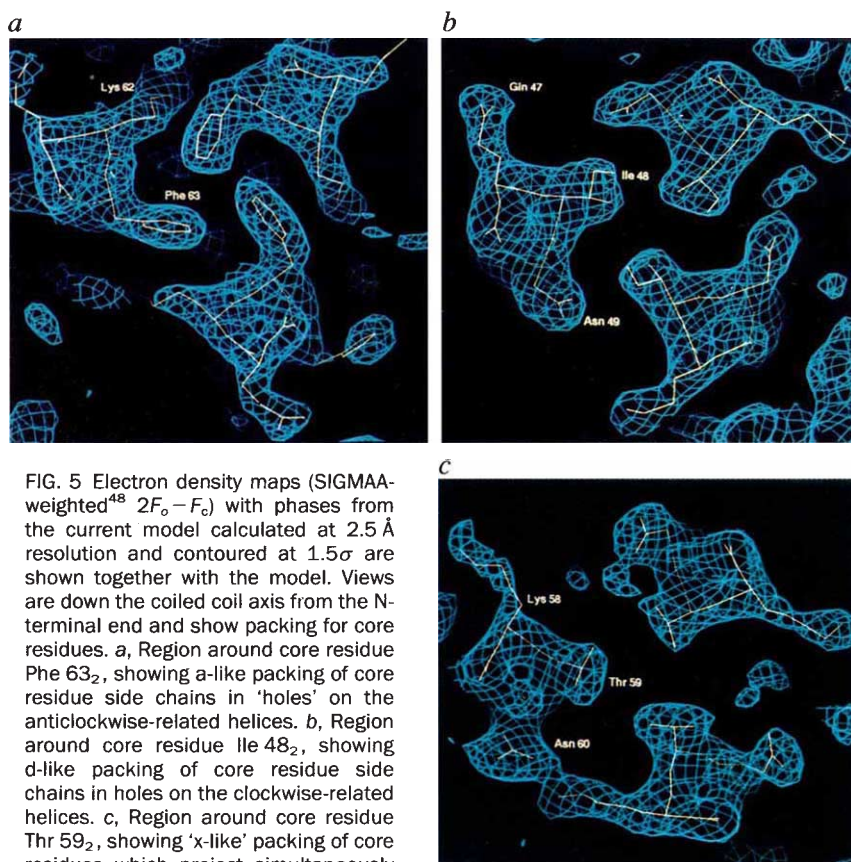


FIG. 5 Electron density maps (SIGMAA-weighted⁴⁸ $2F_o - F_c$) with phases from the current model calculated at 2.5 Å resolution and contoured at 1.5σ are shown together with the model. Views are down the coiled coil axis from the N-terminal end and show packing for core residues. *a*, Region around core residue Phe 63₂, showing a-like packing of core residue side chains in 'holes' on the anticlockwise-related helices. *b*, Region around core residue Ile 48₂, showing d-like packing of core residue side chains in holes on the clockwise-related helices. *c*, Region around core residue Thr 59₂, showing 'x-like' packing of core residues which project simultaneously towards the threefold axis. Core residues are unambiguously a-like or d-like except in three places (Fig. 3a). Two x-like residues are seen near the ends of the coiled coil. The third, shown here, lies in the midst of an otherwise unbroken 3–4 repeat extending from residue 45₂ to 87₂; the spacing for core residues surrounding Thr 59₂ is 3–4–4–3 (Fig. 3a). Thus, it serves as the crossover point between two different but related heptad registers, thereby maximizing the hydrophobicity of the core residues. Another register shift occurs at Leu 91₂, a d-like residue.

the buried location of the N-terminal fusion peptide. The fusion peptide, then, is apparently partially or fully released in the transition to the fusion-active state. A second group (yellow) identifies residues that appear to stabilize the short helix A and the loop B against the long helix CD, suggesting that the transition involves rearrangements of helix A and loop B, perhaps as the residues preceding C are recruited into the coiled coil. The third group (green) are residues that appear to stabilize contacts between HA₁ and B and CD of HA₂, consistent with the proposal that these contacts break as the B loop refolds into a helical conformation and the middle of helix CD refolds into a loop (Fig. 3).

Mutations in the interface between adjacent HA₁ subunits form a fourth group (white, Fig. 4c). No new insights about this group are available directly from the TBHA₂ structure, from which HA₁ residues 28₁–328₁ have been proteolytically removed. As previously discussed^{6,31}, these mutations suggest that the trimeric assembly of HA₁ globular domains rearranges or dissociates at low pH. Movement of the globular domains is probably accompanied by the refolding of the B loop into a helical conformation, as the B loop is the site against which HA₁ is docked in BHA. The HA₁ globular domains themselves apparently retain a native-like structure as they interact with many anti-BHA antibodies^{22,23,26}, bind sialic acid receptor analogues with the same affinity as BHA³², and display an appropriate circular dichroism spectrum¹⁸.

Electron microscopy

The dimensions and shape of TBHA₂ revealed by the X-ray structure agree well with electron micrographs (EMs) of the

molecule. In particular, EM images show molecules 105 ± 7 Å in length and many display a knob at one end, presumably corresponding to the bulky end seen in the X-ray structure¹⁵ (bottom, Fig. 2d). In addition, EM images of an earlier soluble thermolysin digestion product, identical to TBHA₂ except that the predominant HA₂ N terminus is residue 24₂ rather than 38₂, show molecules about 130 Å long, again with terminal knobs¹⁵. This observation is consistent with the idea that residues 24₂–37₂ extend from the tip of the coiled coil, possibly as a continuation of the α -helical coiled coil (~ 21 Å length) or in the β -hairpin conformation they adopt in BHA (~ 22 Å length).

Trypsin-digested low-pH BHA, identical to TBHA₂ except that it retains the first 37 residues of the HA₂ chains including the fusion peptides, forms 'rosettes' of 8 ± 1 molecules with distal knobs¹⁵. Again, these knobs are presumably the bulky 'bottom' region seen in TBHA₂, whereas the fusion peptides mediate aggregation at the centre of the rosette. Virus particles or HA-containing liposomes incubated at pH 5 and then digested with trypsin reveal thin spikes projecting 105 ± 10 Å from the virus membrane or liposome surface²⁴. These spikes have membrane-distal knobs. One model which explains this observation is that both the fusion peptides and the C-terminal transmembrane segments are inserted into the virus or liposomal membrane, causing inversion of the main body of the trimer relative to the membrane. This model is possible given the structure of TBHA₂, inasmuch as the apparently disordered C-terminal region of the HA₂ chains could, in extended conformations, span the distance to the membrane whatever the over-

all orientation of the trimer. Moreover, it provides a possible explanation for the observation that virus becomes inactivated for fusion when incubated at fusion pH in the absence of target membrane³³: namely, that the fusion peptides insert non-productively into the viral membrane.

The removal of 28₁–328₁ and both termini of HA₂ in producing TBHA₂ might have removed residues which stabilize a different conformation, or might have generated additional degrees of freedom unavailable in the intact molecule. While the TBHA₂ structure is consistent with the mutagenic analysis and EM data discussed above, further experiments are required to assess the possibility that, for example, disorder at the chain termini is a result of proteolytic cleavage.

Implications for membrane fusion

The TBHA₂ structure suggests that the fusion-pH-induced conformational change delivers the fusion peptides at least 100 Å towards the target membrane (Fig. 3e), consistent with the expectation that HA might form a bridge between the viral and cellular membranes^{1,11,19,33}. We cannot rule out, however, the possibility that the main body of the molecule inverts to place the fusion peptide in the viral membrane¹⁵. Such inverted molecules might also be observed in the post-fusion state, wherein the transmembrane anchor segments and the fusion peptides may be associated with the same (fused) membrane³⁴. Thus, the low-pH-induced conformation observed here could, in one orientation, mediate membrane fusion and, in another, lead to inactivation.

The increased flexibility inferred from the apparent disorder in the C-terminal region of TBHA₂ suggests that the extra-

membrane domain of HA at the pH of fusion might be able to bend with respect to the plane of the membrane. Such bending could allow the opposed membranes to approach more closely than the length of TBHA₂. In addition, increased orientational flexibility of the HA extramembrane domain might facilitate higher-order assembly of trimers, as proposed to be necessary

for HA-mediated fusion³⁵, perhaps in the formation of a fusion pore structure³⁶. Studies to date have not ruled out the possibility that other parts of the polypeptide, such as the membrane-proximal region inferred to be disordered from the TBHA₂ structure, are directly active in membrane fusion, as suggested for the C-terminal anchor region³⁷. □

Received 31 May; accepted 13 July 1994.

1. Wiley, D. C. & Skehel, J. J. *A. Rev. Biochem.* **56**, 365–394 (1987).
2. Maeda, T. & Ohnishi, S. *FEBS Lett.* **122**, 283–287 (1980).
3. Huang, R. T. C., Rott, R. & Klenk, H.-D. *Virology* **110**, 243–247 (1981).
4. White, J., Matlin, K. & Helenius, A. *J. Cell Biol.* **89**, 674–679 (1982).
5. White, J. M. & Wilson, I. A. *J. Cell Biol.* **105**, 2887–2896 (1987).
6. Godley, L. et al. *Cell* **68**, 635–645 (1992).
7. Brand, C. M. & Skehel, J. J. *Nature new Biol.* **238**, 145–147 (1972).
8. Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 93–97 (1975).
9. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366–373 (1981).
10. Weis, W. I., Brünger, A. T., Skehel, J. J. & Wiley, D. C. *J. molec. Biol.* **212**, 737–761 (1990).
11. Skehel, J. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **79**, 968–972 (1982).
12. Gething, M.-J., Doms, R. W., York, D. & White, J. J. *Cell Biol.* **102**, 11–23 (1986).
13. Daniels, R. S. et al. in *The Origin of Pandemic Influenza Viruses* (eds Laver, G. W. & Chu, C. M.) 1–7 (Elsevier-North Holland, New York, 1985).
14. Doms, R. W., Helenius, A. & White, J. J. *J. biol. Chem.* **260**, 2973–2981 (1985).
15. Ruigrok, R. W. H. et al. *J. gen. Virol.* **69**, 2785–2795 (1988).
16. Bullough, P. A. et al. *J. molec. Biol.* **236**, 1262–1265 (1994).
17. Brünger, A. T. *Nature* **355**, 472–475 (1992).
18. Wharton, S. A. et al. *J. biol. Chem.* **263**, 4474–4480 (1988).
19. Carr, C. M. & Kim, P. S. *Cell* **73**, 823–832 (1993).
20. Ward, C. W. & Dopheide, T. A. *Aust. J. biol. Sci.* **33**, 441–447 (1980).
21. Graves, P. N., Schulman, J. L., Young, J. F. & Palese, P. *Virology* **126**, 106–116 (1983).
22. Webster, R. G., Brown, L. E. & Jackson, D. C. *Virology* **126**, 587–599 (1983).
23. Daniels, R. S. et al. in *The Origin of Pandemic Influenza Viruses* (eds Laver, G. W. & Chu, C. M.) 8–18 (Elsevier-North Holland, New York, 1985).
24. Ruigrok, R. W. H. et al. *EMBO J.* **5**, 41–49 (1986).
25. Daniels, R. S. et al. *EMBO J.* **6**, 1459–1465 (1987).
26. Yewdell, J. W., Gerhard, W. & Bachi, T. J. *Virology* **155**, 484–497 (1986).
27. Ruigrok, R. W. H. et al. *Virology* **155**, 484–497 (1986).
28. Goldsmith, E. J. & Mottonen, J. *Structure* **2**, 241–244 (1994).
29. Crick, F. H. C. *Acta crystallogr.* **6**, 689–697 (1953).
30. McLachlan, A. D. & Karn, J. J. *J. molec. Biol.* **164**, 605–626 (1983).
31. Daniels, R. S. et al. *Cell* **40**, 431–439 (1985).
32. Sauter, N. K. et al. *Biochemistry* **31**, 9609–9621 (1992).

33. Stegmann, T. & Helenius, A. in *Viral Fusion Mechanisms* (ed. Bentz, J.) 89–111 (CRC Press, Boca Raton, Florida, 1993).
34. Söllner, T., Bennett, M. K., Whiteheart, S. W., Scheller, R. H. & Rothman, J. E. *Cell* **75**, 409–418 (1993).
35. Ellens, H., Bentz, J., Mason, D., Zhang, F. & White, J. M. *Biochemistry* **29**, 9697–9707 (1990).
36. Spruce, A. E., Iwata, A., White, J. M. & Almers, W. *Nature* **342**, 555–558 (1989).
37. Kemble, G. W., Danielli, T. & White, J. M. *Cell* **76**, 383–391 (1994).
38. Blum, M., Metcalf, P., Harrison, S. C. & Wiley, D. C. *J. appl. Crystallogr.* **20**, 235–242 (1987).
39. Leslie, A. G. W. in *CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, No. 26 (Daresbury Laboratory, Warrington, 1992).
40. *CCP4, A Suite of Programs for Protein Crystallography* (SERC (UK) Collaborative Computing Project No. 4, Daresbury Laboratory, Warrington, 1979).
41. Terwilliger, T. C. & Eisenberg, D. *Acta crystallogr.* **A39**, 813–817 (1983).
42. Metcalf, P., Blum, M., Freymann, D., Turner, M. & Wiley, D. C. *Nature* **325**, 84–86 (1987).
43. Bricogne, G. *Acta crystallogr.* **A32**, 832–847 (1976).
44. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110–119 (1991).
45. Brünger, A. T. *X-PLOR (Version 3.1): A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).
46. Lüthy, R., Bowie, J. U. & Eisenberg, D. *Nature* **356**, 83–85 (1992).
47. Colovos, C. & Yeates, T. O. *Prot. Sci.* **2**, 1511–1519 (1993).
48. Read, R. J. *Acta crystallogr.* **A42**, 140–149 (1986).
49. Kraulis, P. J. *J. appl. Crystallogr.* **24**, 924–950 (1991).
50. Carson, M. J. *J. molec. Graphics* **5**, 103–106 (1987).

ACKNOWLEDGEMENTS. P.A.B. and F.M.H. contributed equally to this work. We dedicate this structure determination to Max Perutz, FRS, a founder of our science, on the year of his eightieth birthday. We thank R. Daniels, R. Ruigrok, A. Trehanne and J. Newman for earlier contributions; E. Collins, S. Garman, P. Rosenthal, L. Stern, S. Watowich, and the staff of the Cornell High Energy Synchrotron source MacCHESS for help with data collection; D. Stevens and A. Douglas for technical assistance; and our colleagues for advice and encouragement. F.M.H. was supported by a Helen Hay Whitney Foundation fellowship and the HHMI. P.A.B. was supported by the NIH. This work was funded by the NIH and the MRC. D.C.W. is an investigator with the Howard Hughes Medical Institute. Coordinates are available from the authors (email: hughson@XTAL.Ø.harvard.edu).

LETTERS TO NATURE

Measurement of the microwave background temperature at a redshift of 1.776

A. Songaila*, L. L. Cowie*, S. Vogt†, M. Keane†, A. M. Wolfe‡, E. M. Hu*, A. L. Oren‡, D.-R. Tytler‡ & K. M. Lanzetta§

* Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

† Lick Observatory, University of California, Santa Cruz, California 95064, USA

‡ Department of Astrophysics and Space Science, University of California, San Diego, La Jolla, California 92093, USA

§ Astronomy Program, Department of Earth and Space Science, State University of New York, Stony Brook, New York 11794, USA

HOT Big Bang cosmology predicts that the temperature of the cosmic microwave background radiation will increase linearly with increasing redshift to early in the history of the Universe. The local background temperature (2.7 K) is known very accurately from direct measurements^{1–3}, but other techniques must be used to estimate it at non-zero redshifts. One way is to determine the excitation of atomic transitions in absorbing clouds along the lines-of-sight to distant quasars⁴. When the transitions are in equilibrium with the microwave background radiation, the radiation will populate the fine-structure levels of the ground states of certain atoms,

and the relative populations of the levels can be used to calculate its temperature. Here we report the detection of absorption from the first fine-structure level of neutral carbon atoms in a cloud at a redshift of 1.776, towards the quasar Q1331 + 170. The population ratio yields a temperature of 7.4 ± 0.8 K, assuming that no other significant sources of excitation are present. This agrees with the theoretical prediction of 7.58 K.

The cosmic microwave background radiation (CMBR) will populate excited levels of atomic and molecular species when the energy separations involved are not too different from the CMBR peak frequency. The first measurement of the local CMBR temperature was in fact made using this method⁵ with fine structure lines in the cyanogen (CN) molecule, although it was not recognised as such until after Penzias and Wilson identified the CMBR⁶. Cyanogen excitation can now be used to measure T_{CMBR} very precisely. Roth *et al.*⁷, who measured the rotational excitation of CN toward five Galactic stars and carefully corrected for local sources of excitation, found a value of T_{CMBR} at 2.64 mm of $2.729^{(+0.023)}_{(-0.031)}$ K, in agreement with the COBE result² of 2.726 ± 0.010 K.

Bahcall and Wolf⁸ first suggested that the method could be extended to high redshift, where T_{CMBR} could be presumed to be larger, using atomic fine-structure transitions in absorbing clouds toward high-redshift quasars. Useful transitions for this purpose include those of C⁰, C⁺ and N⁺, with C⁰ being particularly well suited^{4,8}. This measurement has been attempted several times, but has generally been limited by the resolution and signal-to-noise available in reasonable exposure times at the very faint magnitudes involved, and, in the case of C⁰, the intrinsic weakness of the line. C⁺ and N⁺ have strong lines, common in the spectra of quasars, but their relatively high fine-structure

splitting makes them rather insensitive probes of the temperature. Furthermore the N^+ line lies in the region of quasar spectra confused by the Ly- α 'forest' of intergalactic neutral hydrogen lines. Nevertheless, Bahcall, Joss and Lynds⁹ found an upper limit of $T_{\text{CMBR}} < 45$ K at redshift $z=2.309$ toward the quasar PHL957 from C^+ absorption. The most recent measurement of Songaila *et al.*¹⁰, with very high signal-to-noise observations from the Keck telescope (Mauna Kea, Hawaii), achieved a limit of $T_{\text{CMBR}} < 13.5$ K at $z=2.909$ towards the quasar Q0636+680 based on upper limits to C^+ fine structure. C^0 excitation is inherently more sensitive, because of the small splitting of 16.4 cm^{-1} of the $J=0$ and $J=1$ levels of the ground state, but C^0 absorption is very rare in quasar spectra. The best measurement to date with C^0 is that of Meyer *et al.*¹¹ in the spectrum of Q1331+170. Their upper limit on the column density of carbon atoms in the $J=1$ level ($C \text{ I}^*$) implies that $T_{\text{CMBR}} < 16$ K at $z=1.776$. We now report the first detection of $C \text{ I}^*$ at high redshift, toward this same system in Q1331+170.

Our measurements were made with the HIRES high-resolution spectrograph¹² at the Keck 10-m telescope on 16–18 April 1994 (UT). Thirteen 60-min exposures of the quasar Q1331+170 were used, taken with a measured resolution $R=47,000$, which corresponds to an instrumental 'b-parameter' of 3.6 km sec^{-1} . ($b \equiv (2kT/m)^{1/2}$, where m is the atomic mass and T the temperature.) The data were reduced as described in ref. 10 and the resulting spectra near the C^0 multiplets at $1,656 \text{ \AA}$ and $1,560 \text{ \AA}$ are shown in Fig. 1. An energy level diagram showing the transitions in the relevant multiplets is given in Fig. 2. The $J=0$ ($C \text{ I}$) and $J=1$ ($C \text{ I}^*$) absorption occurs in two main velocity components with (vacuum heliocentric) redshifts of 1.77638 and 1.77654, as well as some weak higher-velocity absorption, giving a weighted mean redshift of 1.77644 ± 0.00002 , in good agreement with the 21-cm redshift¹³ of 1.77642 ± 0.00002 . We fitted Voigt profiles simultaneously to all the observable transitions in the two multiplets to obtain the column densities in the $C \text{ I}$ and $C \text{ I}^*$ lines. The fitting parameters of the two main components are detailed in Table 1. Absorption from the $J=2$ level is not detected with a 1σ upper limit of $5 \times 10^{11} \text{ cm}^{-2}$. Component 1 is strongest, both in the ground state, where the C^0 column density, $N=1.25 \times 10^{13} \text{ cm}^{-2}$, and in the $J=1$ state, where $N=3.9 \times 10^{12} \text{ cm}^{-2}$. Component 2 is weaker in the ground state, with $N=7.2 \times 10^{12} \text{ cm}^{-2}$, but very weak in the $J=1$ state. In the individual lines it is detected only at the $(1-2)\sigma$ level, and in

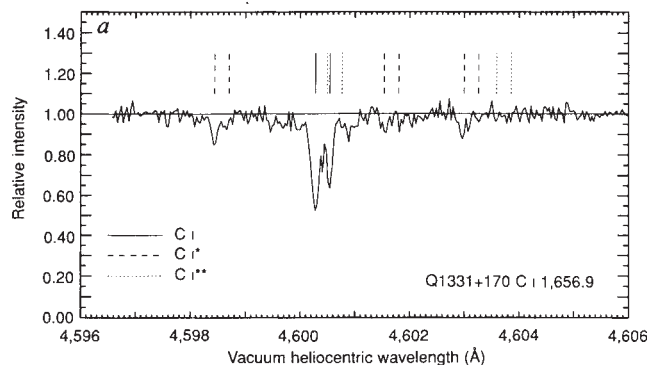


FIG. 1 Portions of the spectrum of the quasar Q1331+170: a, near $C \text{ I}$ multiplet 2 at rest wavelength $1,656.9 \text{ \AA}$ (orders 77 and 78); b, near $C \text{ I}$ multiplet 3 at rest wavelength $1,560.3 \text{ \AA}$ (order 82) in the $z=1.77642$ absorption-line system. Two main velocity components are seen, at vacuum heliocentric redshifts of 1.77638 and 1.77654, as well as some weak higher-velocity absorption. Vertical solid lines show the positions of the ground-state transitions of components 1 and 2, dashed lines the $C \text{ I}^*$, and dotted lines the $C \text{ I}^{**}$ fine-structure lines. Individual positions for the $C \text{ I}^*$ lines in multiplet 3 are given, although these are

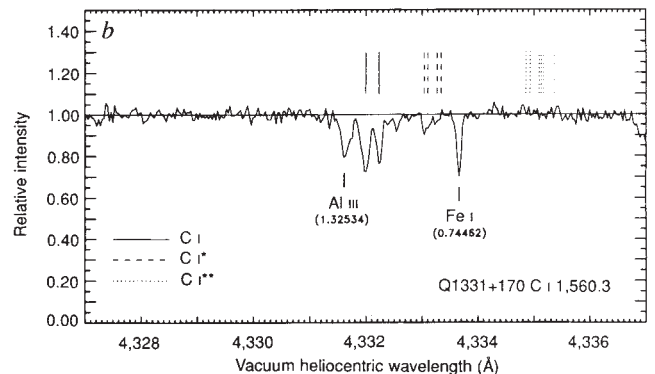
	Component 1	Component 2
Redshift	1.77638	1.77654
Multiplet 2		
$b \text{ (km s}^{-1}\text{)}$	6.4	2.1
$N(\times 10^{12} \text{ cm}^{-2})^*$:		
$J=0, \lambda=1,656.928 \text{ \AA}$	12.6 ± 0.7	6.7 ± 0.5
$J=1, \lambda=1,656.267 \text{ \AA}$	5.5 ± 1.0	0.9 ± 0.8
$J=1, \lambda=1,657.379 \text{ \AA}$	5.6 ± 2.7	1.9 ± 1.6
$J=1, \lambda=1,657.907 \text{ \AA}$	6.4 ± 2.1	1.0 ± 1.2
$\langle J=1 \rangle$ (weighted mean)	5.6 ± 0.9	1.1 ± 0.6
$N(J=1)/N(J=0)$	$0.44^{(+0.08)}_{(-0.07)}$	$0.16^{(+0.09)}_{(-0.09)}$
$T_{\text{excitation}} \text{ (K)}$	$12.3^{(+1.2)}_{(-1.1)}$	$8.0^{(+1.5)}_{(-1.5)}$
Multiplet 3		
$b \text{ (km s}^{-1}\text{)}$	6.4	2.6
$N(\times 10^{12} \text{ cm}^{-2})^*$:		
$J=0, \lambda=1,560.309 \text{ \AA}$	12.5 ± 0.5	7.5 ± 0.4
$J=1^\dagger, \lambda=1,560.695 \text{ \AA}$	3.4 ± 0.5	0.75 ± 0.4
$N(J=1)/N(J=0)$	$0.27^{(+0.04)}_{(-0.04)}$	$0.10^{(+0.05)}_{(-0.05)}$
$T_{\text{excitation}} \text{ (K)}$	$9.8^{(+0.8)}_{(-0.6)}$	$7.0^{(+1.2)}_{(-0.9)}$
Weighted mean of multiplets		
$N(J=0) (\times 10^{12} \text{ cm}^{-2})$	12.5 ± 0.4	7.2 ± 0.3
$N(J=1) (\times 10^{12} \text{ cm}^{-2})$	3.9 ± 0.4	0.9 ± 0.3
$T_{\text{excitation}} \text{ (K)}$	10.4 ± 0.5	7.4 ± 0.8

Symbols used: N is the column density and b the width of the line ($\equiv (2kT/m)^{1/2}$ where m is the atomic mass and T the temperature). $T_{\text{excitation}}$ is the excitation temperature inferred from the relative populations of the $J=1$ and $J=0$ levels of the ground state.

* Errors are based on random fitting to 20 surrounding positions. Dropping the continuum level by 1% (an extreme case) would decrease the ratio of component 2 an amount comparable to the 1σ error.

† The average of the $J=1$ lines at $1,560.6832 \text{ \AA}$ and $1,560.7079 \text{ \AA}$, with combined oscillator strength $f=0.081$. The $C \text{ I}^*$ lines are blended at our resolution. All rest wavelengths and oscillator strengths are from Morton¹⁹.

the weighted mean it is present only at the 3σ level, with $N=(9 \pm 3) \times 10^{11} \text{ cm}^{-2}$. Component 2 is very narrow and is only marginally resolved at our resolution, with $b \leq 2.6 \text{ km sec}^{-1}$, which corresponds to an upper limit on the cloud kinetic temperature of $6,500 \text{ K}$. Multiplet equivalent-width ratios show that $b \geq 1 \text{ km sec}^{-1}$, and that the lines are unsaturated. The column



blended at our resolution of $R=47,000$. The neighbourhood of multiplet 3 contains lines of Al III ($1,862.790 \text{ \AA}$) at $z=1.32534$ and Fe I ($2,484.021 \text{ \AA}$) at $z=0.74462$ whose positions are marked. The $z=0.74462$ system contains many strong lines and is well known²⁰. The $z=1.32534$ system has not been reported before, but is confirmed by the identification of the other line of the Al III doublet ($1,854.716 \text{ \AA}$) at $4,312.0 \text{ \AA}$ in order 82, by Mg II ($2,795.528 \text{ \AA}$) absorption at $6,501.1 \text{ \AA}$ and weak Fe II absorption ($2,344.214$, $2,382.765$ and $2,600.173 \text{ \AA}$) at $5,448.4$, $5,539.5$ and $6,045.1 \text{ \AA}$, respectively.

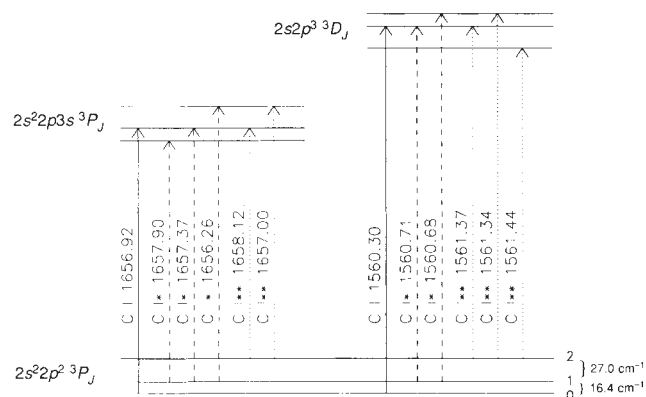


FIG. 2 Schematic energy-level diagram showing the ground state and fine-structure transitions of the two C^0 multiplets at 1,656 and 1,560 Å. The fine-structure energy splittings of the ground state are also shown (16.4 cm^{-1} , $J=1-0$; 27.0 cm^{-1} , $J=2-1$).

densities are therefore insensitive to the b values. If there were unseen narrow saturated components that do not dominate the equivalent width, the ground-state column density would be larger and our measured excitation temperature would be lowered. Our measured ratios of population in the $J=1$ and $J=0$ levels, $N(J=1)/N(J=0)=0.31 \pm 0.03$ in component 1 and 0.125 ± 0.04 in component 2, imply excitation temperatures of $10.4 \pm 0.5 \text{ K}$ in component 1 and $7.4 \pm 0.8 \text{ K}$ in component 2. The predicted value of the temperature at this redshift is $T_{\text{CMBR}} = 7.58 \text{ K}$.

In the interstellar medium (ISM) of the Galaxy, because T_{CMBR} is much smaller than the C^0 fine-structure separation, the $J=1$ level is populated predominantly by collisions with neutral hydrogen¹⁴. A small additional contribution is provided by optical pumping by the ultraviolet radiation field, but direct excitation by the CMBR is generally negligible¹⁴. However, measured values of $C\text{ I}^*/C\text{ I}$ in the Galactic ISM can be as high as 1.0, which encompasses the values measured in both of the $C\text{ I}$ components in Q1331+170. In contrast, excitation by the CMBR is substantial at $z=1.78$, and optical pumping is not important unless the ultraviolet radiation field is more than an order of magnitude larger than in the Galaxy¹⁵. The $J=0-1$ absorption from a background radiation field of $T=7.58 \text{ K}$ is $1.1 \times 10^{-8} \text{ s}^{-1}$ (ref. 15), compared to a $J=0-1$ excitation rate of $3.8 \times 10^{-10} n_{\text{H}} \text{ s}^{-1}$ at a cloud kinetic temperature of 100 K and $6.6 \times 10^{-10} n_{\text{H}} \text{ s}^{-1}$ at 1,000 K from collisions with neutral hydrogen¹⁶. (n_{H} is the density of neutral hydrogen in atoms per cm^3 , and the two temperatures are representative of a reasonable

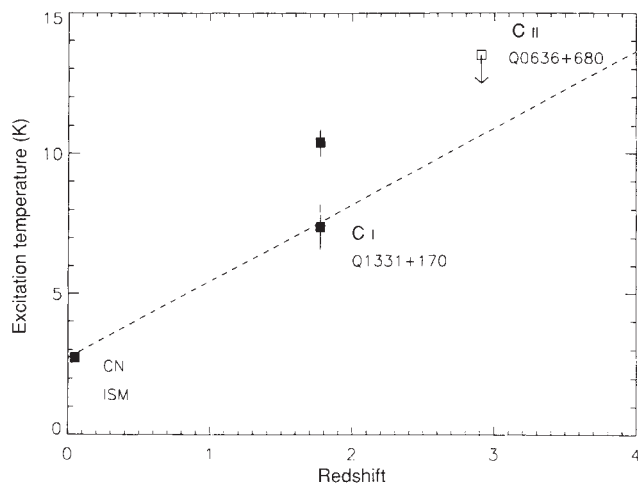


FIG. 3 Measurements of the excitation temperature at various redshifts. The filled square at $z=0$ (offset to $z=0.05$ for clarity) shows the result of Roth, Meyer and Hawkins⁷ determined from measurements of interstellar medium (ISM) CN. The upper limit at $z=2.909$ (hollow square) was derived in ref. 10 using $C\text{ II}$ in an absorption system toward Q0636+680. The present measurements at $z=1.776$ are shown as filled squares. We show 1σ error bars (which are smaller than the plotting symbol for the $z=0$ case). The dashed line is the prediction $T_{\text{CMBR}}(z) = T_{\text{CMBR}}(0)(1+z)$.

range of kinetic temperatures in a cloud containing $C\text{ I}$ (ref. 14). So, given that the CMBR is at its expected temperature, we find 1σ upper limits of $n_{\text{H}}=7 \text{ cm}^{-3}$ for $T=100 \text{ K}$, and 4 cm^{-3} for $T=1,000 \text{ K}$, and the pressure in the system is low compared to that in our own interstellar medium, but not unreasonably so¹⁴. Unfortunately the excitation of other fine-structure lines, such as those of Si^+ and C^+ , does not usefully constrain n_{H} at these levels.

Unless the issue of local excitation can be more completely addressed, our T_{CMBR} value must be considered to be an upper limit, although one strikingly close to the standard prediction; if it could be shown that significant other sources of excitation were present, the measured value would be less, possibly leading to a conflict with the Big Bang expectation. In terms of $T_{\text{CMBR}}(z) = T_{\text{CMBR}}(0)(1+z)^{\alpha}$, we find $\alpha = 0.98 \pm 0.11$ (Fig. 3). Of course, the redshift of Q1331+170 is quite low. It would be very gratifying to find an appropriate C^0 absorber at higher redshift to improve existing, much less sensitive, limits from C^+ . Although the modern steady-state model¹⁷ predicts $T_{\text{CMBR}} \propto (1+z)$ until $z \sim 6$, other theories¹⁸ make strong claims for a value of T_{CMBR} well below the standard prediction by $z \sim 3$, which we could check by such a measurement. □

Received 18 May; accepted 8 July 1994.

- Mather, J. C. et al. *Astrophys. J.* **354**, L37–L40 (1990).
- Mather, J. C. et al. *Astrophys. J.* **420**, 439–444 (1994).
- Smoot, G. F. et al. *Astrophys. J.* **396**, L1–L5 (1992).
- Bahcall, J. N. & Wolf, R. A. *Astrophys. J.* **152**, 701–729 (1968).
- Thaddeus, P. A. *Rev. Astr. Astrophys.* **10**, 305–332 (1972).
- Penzias, A. A. & Wilson, R. W. *Astrophys. J.* **142**, 419–421 (1965).
- Roth, K. C., Meyer, D. M. & Hawkins, I. *Astrophys. J.* **413**, L67–L71 (1993).
- Khersonskii, V. K. *Soviet Astr.* **25**, 134 (1981).
- Bahcall, J. N., Joss, P. C. & Lynds, R. *Astrophys. J.* **182**, L95–L98 (1973).
- Songaila, A., Cowie, L. L., Hogan, C. & Rugers, M. *Nature* **368**, 599–604 (1994).
- Meyer, D. M., Black, J. H., Chaffee, F. H., Foltz, C. & York, D. G. *Astrophys. J.* **308**, L37–L41 (1986).
- Vogt, S. S. et al. *S.P.I.E.* (in the press).

- Wolfe, A. M. & Davis, M. M. *Astr. J.* **84**, 699–706 (1979).
- Jenkins, E. B. & Shaya, E. J. *Astrophys. J.* **231**, 55–72 (1979).
- Chaffee, F. H., Black, J. H. & Foltz, C. B. *Astrophys. J.* **335**, 584–592 (1988).
- Launay, J. M. & Roueff, E. *Astr. Astrophys.* **56**, 289–292 (1977).
- Hoyle, F., Burbidge, G. & Narlikar, J. V. *Astrophys. J.* **410**, 437–457 (1993).
- Phillips, P. R. *Mon. Not. R. astr. Soc.* (in the press).
- Morton, D. C. *Astrophys. J. Suppl. Ser.* **77**, 119–202 (1991).
- Junkkarinen, V., Hewitt, A. & Burbidge, G. *Astrophys. J. Suppl. Ser.* **77**, 203–254 (1991).

ACKNOWLEDGEMENTS. We are grateful to T. Bida, P. Gillingham, J. Aycock, T. Chelminiak and W. Wack for their extensive help in obtaining the observations. We particularly thank B. Savage for his comments on an early draft of the Letter. A.S., L.L.C. and E.M.H. are Visiting Astronomers at the W. M. Keck Observatory, jointly operated by the California Institute of Technology and the University of California. This work was supported at the University of Hawaii by the State of Hawaii and by NASA.

A superluminal source in the Galaxy

I. F. Mirabel* & L. F. Rodríguez†

* Service d'Astrophysique/CEA/DSM/DAPNIA, Centre d'Etudes de Saclay, F-91191 Gif-sur-Yvette, France

† National Radio Astronomy Observatory, PO Box 0, Socorro, New Mexico 87801, USA, and Instituto de Astronomía, UNAM, Mexico

APPARENT velocities greater than the speed of light (superluminal motion) have been inferred for radio-emitting components in a number of distant quasars and active galactic nuclei¹. These components move away from the central sources (generally thought to be super-massive black holes) at rates that seem to imply velocities greater than c . The accepted explanation is that clouds of plasma are ejected in opposite directions from the central source at speeds close to (but less than) that of light, and that relativistic effects lead to the apparent superluminal motion². But the extreme distance of the objects observed so far introduces many uncertainties into this interpretation³. Here we present observations of the first apparent superluminal motion ever detected in a source within our own Galaxy. The optical, infrared and X-ray properties^{4,5} of the counterpart suggest that the source is either a neutron star or a black hole that is ejecting matter in a process similar to, but on a smaller scale than that seen in quasars. Because of its relative proximity, this superluminal microquasar may offer the best opportunity to gain a general understanding of relativistic ejections seen elsewhere in the Universe.

In the course of a general study of radio counterparts of γ -ray sources we observed with the Very Large Array a remarkable ejection event in the transient source GRS1915+105 (ref. 6). After the observation of strong radio outbursts^{7,8}, we have followed the large proper motions of a pair of bright radio condensations emerging in opposite directions from a compact core. Figure 1 shows maps of GRS1915+105 at a wavelength of 3.5 cm for seven epochs of observation between 27 March and 30 April 1994. These composite maps exhibit a pair of bright radio condensations (see Table 1) emanating from a flux-variable, compact source. The angular resolution of the observations reported here was 0.2 arcsec, and the positions were determined by absolute astrometry with accuracies of 0.02 arcsec.

In Fig. 2 are plotted for each condensation the angular displacements from the stationary core as a function of time. This Figure shows that the components move away with different apparent speeds, and that both were expelled from the compact source on 19 March 1994 at 20:00 $\pm$ 5 UT, when the source appeared brighter than usual in the X-rays⁹.

In the following we model this object on the basis of the anti-parallel ejection of a pair of condensations moving at speed $\beta(\beta = v/c$, with v being their velocity and c the speed of light),

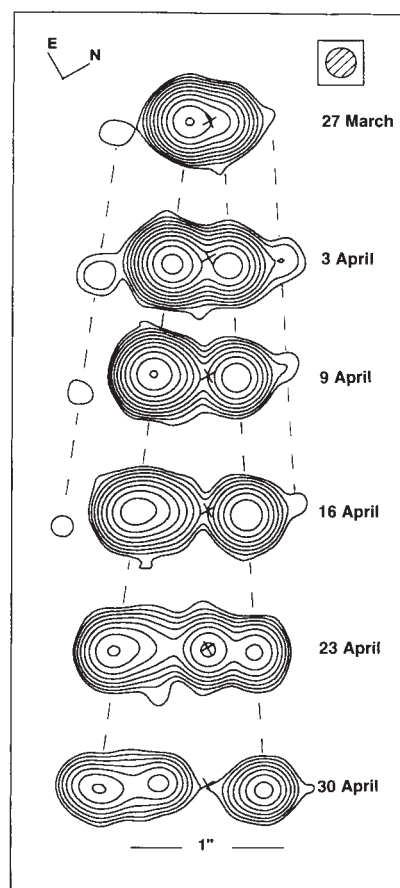


FIG. 1 Pair of bright radio condensations moving away from the high-energy source GRS1915+105. These uniform-weight, self-calibrated VLA maps were made at wavelengths of 3.5 cm for the 1994 epochs shown on the right side of the Figure. The vertical separation between the individual maps is proportional to the interval of time between the observations. Note in the first four epochs the presence of a fainter pair of condensations moving ahead of the bright ones at about the same speed. On 23 April the core flared again, and on 30 April a new southern component appeared. The stationary position of the core at right ascension $\alpha(1950) = 19^{\text{h}} 12^{\text{m}} 49.966^{\text{s}}$; declination $\delta(1950) = 10^{\circ} 51' 26.73'' (\pm 0.02'')$, which is indicated with a small cross, was determined by absolute astrometry from maps made at times when the core was observed flaring (26 February, 3 March, 18 March and 23 April 1994). The half-power beam width of 0.2 arcsec is shown in the top-right corner. The maps have been rotated 60° clockwise for easier display and the actual orientations of the north and east are also shown in the top-left corner. Contours are 1, 2, 4, 8, 16, 32, 64, 128, 256, 512 and 1,024 times 0.2 mJy per beam for all epochs except for 27 March, where the contour levels are the above in multiples of 0.6 mJy per beam.

TABLE 1 Evolution of the radio condensations at 3.5-cm wavelengths

Date (1994)	Hour (UT)	South component					North component				
		Dist. (")	PA (°)	Flux (mJy)	Size (mas)	p.a. (°)	Dist. (")	PA (°)	Flux (mJy)	Size (mas)	p.a. (°)
24 Mar	14:75	0.08	155	655	60 × 20	167					
27 Mar	16:33	0.14	146	329	50 × 30	121	0.07	330	102	50 × 30	313
03 Apr	15:58	0.24	157	150	90 × 30	140	0.13	313	47	100 × 30	328
09 Apr	15:92	0.36	147	124	110 × 50	148	0.18	331	25	60 × 50	319
16 Apr	13:00	0.45	149	63	210 × 60	156	0.25	326	27	60 × 40	314
23 Apr	12:58	0.59	150	42	230 × 70	160	0.31	327	16		
30 Apr	12:43	0.72	150	36	570 × 80	154	0.37	327	20	90 × 50	340

To derive the spectral indices at different stages of evolution on 24 March and 16 April we also carried out observations at wavelengths of 2 and 6 cm. On 24 March, 3.8 d after the ejection, when the pair of condensations still could not be resolved, the spectral index ($S \propto \nu^{\alpha}$) was -0.49 ± 0.01 . On 16 April, when the two condensations had moved apart, both had spectral index $\alpha = -0.84 \pm 0.03$. PA is the position angle of the peak emission with respect to the stationary core, and p.a. of the major axis. Flux density errors are $\sim 5\%$.

TABLE 2 Multi-wavelength properties of GRS1915+105 and SS433

Source	μ arcsec yr ⁻¹	Dist. (kpc)	Veloc. ejecta	Precession of axis	Emission (keV)	L_x (erg s ⁻¹)	A_v (mag)	M_K (mag)	ΔK (mag)
GRS1915+105	6.4/3.3	12.5	0.92c	Unknown	≤ 220	3×10^{38}	≥ 20	-6.0	0.9
SS433	3.2	5.5	0.26c	164 d	≤ 30	5×10^{35}	7	-5.5	1.1

Symbols used: μ , proper motion; L_x , X-ray luminosity; A_v , absorption at visual wavelengths; M_K , absolute infrared magnitude in the $K=2.2 \mu\text{m}$ band; ΔK , observed variation in the infrared K band.

with the axis of the flow making an angle θ with respect to the line of sight of a distant observer. The apparent proper motions of the approaching and receding condensations, μ_a and μ_r , are given¹⁰ by:

$$\mu_{r,a} = \frac{\beta \sin \theta}{(1 \pm \beta \cos \theta)} \frac{c}{D} \quad (1)$$

where D is the distance from the observer to the source. The asymmetry in the speeds of separation from the compact core between the southern and northern condensations that is seen in Figs 1 and 2 may be explained as the result of the differential time delay between the approaching and receding components², which is expressed by opposite signs in the denominator of equation (1).

Due to relativistic aberration, the ratio of the apparent surface brightnesses (measured at equal angular separations from the core) from twin components moving at high velocities in opposite directions is given¹⁰ by

$$\frac{S_a}{S_r} = \left(\frac{1 + \beta \cos \theta}{1 - \beta \cos \theta} \right)^{k-\alpha} \quad (2)$$

where α is the spectral index of the radio-emitting components. In the case of continuous jets $k=2$, whereas for discrete condensations $k=3$. Because (as we show below) $\beta \cos \theta = 0.323$ and $\alpha = -0.8$ (see legend of Table 1), the flux ratio in the case of twin discrete condensations would be 12, whereas for continuous jets it would be 6. The fluxes listed in Table 1 show that for equal angular separations, the flux ratio between condensations is 8 ± 1 . Therefore the asymmetries, exhibited both in the speeds of separation from the compact core and in the flux ratios, allows us (under the assumption of relativistic motions), to identify the southern component as approaching and the northern as receding.

A relativistic upper limit for the distance can be derived from the measured proper motions of the approaching and receding condensations. Equation (1) can be transformed to the equivalent pair of equations:

$$\beta \cos \theta = \frac{\mu_a - \mu_r}{\mu_a + \mu_r} \quad (3)$$

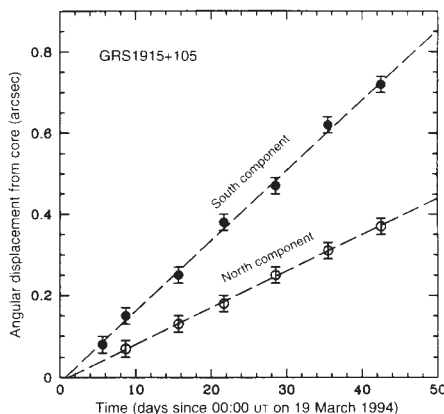


FIG. 2 Angular displacements of the bright radio-emitting components from the compact core as a function of time. The southern condensation (solid circles) appears to move faster than the northern condensation (empty circles). The best fits for the components are, respectively, $\mu_s = 17.6 \pm 0.4 \text{ mas d}^{-1}$ along a direction with position angle $150 \pm 3^\circ$, and $\mu_n = 9.0 \pm 0.1 \text{ mas d}^{-1}$ along a direction with position angle $330 \pm 5^\circ$ (empty circles). Both regression lines converge toward the same point on the time axis. This implies that the major pair of condensations was ejected from the compact object about 19 March 1994, at 20:00 $\pm 5 \text{ ut}$.

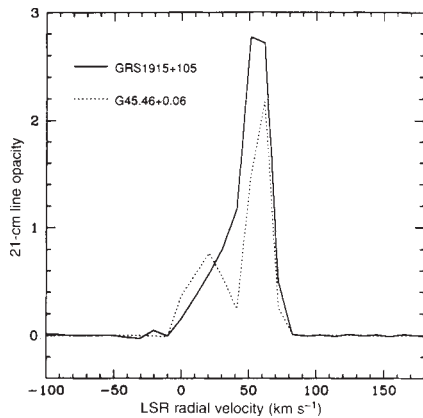


FIG. 3 Opacities of 21-cm line from atomic hydrogen absorption spectra along the lines of sight to GRS1915+105 and to the H II region G45.46+0.06, at a projected distance of 17 arcmin from the high-energy source. These VLA-D observations were carried out on 16 December 1993, when we measured from the high-energy source a continuum flux at 21 cm of 230 mJy. The radial velocity with respect to the Local Standard of Rest (LSR) of $+53.6 \pm 1 \text{ km s}^{-1}$ for G45.46+0.06 obtained from H109 α observations¹¹ places the H II region unambiguously at ~ 8.8 kpc from the Sun (assuming that the Sun is at a Galactocentric distance of 8.5 kpc). From the additional H I absorption observed towards GRS1915+105 (note the significant excess of opacity at $\sim +40 \text{ km s}^{-1}$ from H I at ~ 10 kpc) we derive for the high-energy source a kinematic distance of 12.5 ± 1.5 kpc.

The large kinetic energy of the condensations suggests an acceleration mechanism with very large power. From the source size on 24 March, 5 d after the ejection, we estimate that the ejection event must have lasted ≤ 3 d, requiring a minimum power of $1.2 \times 10^{41} \text{ erg s}^{-1}$ to accelerate the twin ejecta. This is ≥ 400 times the maximum observed steady photon luminosity of GRS1915+105 (ref. 5), which is $3 \times 10^{38} \text{ erg s}^{-1}$. Because on 18 and 21 March, the measured⁹ steady X-ray flux from GRS1915+105 was $\leq 10^{39} \text{ erg s}^{-1}$, a radiation acceleration mechanism would require an increase by $\geq 10^2$ in the source's luminosity some time between 18 and 21 March. We note that the power required for the acceleration of the twin condensations exceeds the luminosity of the soft γ -ray bursts observed in previous years from the same region of the sky¹², which had values of $\sim 10^{40} \text{ erg s}^{-1}$. Because of the large position errors of the γ -ray telescopes, the association of GRS1915+105 with one of the only three known repeaters of soft γ -ray bursts is still uncertain^{4,12–14}.

GRS1915+105 often becomes the most powerful X-ray emitter in the Galaxy. The X-ray light curve⁵ observed since its discovery⁶ in 1992 shows that, for recurrent periods of several months, it is one of the brighter sources of the sky at energies $\geq 20 \text{ keV}$. At a distance of 12.5 kpc, its mean X-ray luminosity during periods of activity is $\sim 3 \times 10^{38} \text{ erg s}^{-1}$, ~ 10 times the mean luminosity of the classic black-hole candidate Cygnus X-1. The spectra^{5,15} of GRS1915+105, with emission up to 220 keV and photon indices that change between -2.5 and -3.0 , indicate that the source is a massive stellar remnant (black hole or neutron star).

This superluminal microquasar appears to be a scaled-up version of the famous stellar source of radio jets SS433 (ref. 16), which was believed to be unique in our Galaxy. For comparison, we list in Table 2 some of the overall properties of these two objects. Although the radio properties of GRS1915+105 may appear reminiscent of SS433, the actual velocity of the ejecta and the X-ray luminosity are much larger in GRS1915+105. In particular, the kinetic energy per unit mass of the ejecta, that is

proportional to $(\gamma - 1)$, is ~ 40 times larger in GRS1915+105. However, similar large ratios of 'kinetic' to photon luminosity are found for both sources. Furthermore, in both sources the kinetic energy of the electrons due to the bulk motion of the ejecta is comparable to the high energy cut-off in the X-ray spectra, suggesting strong coupling via Compton scattering.

The study of Doppler-shifted spectral lines in GRS1915+105, as was done for SS433 (ref. 16), would allow the system to be resolved, leading to an independent determination of the distance by a method that has no precedents in astronomy. The Doppler factors (the ratios of observed to emitted frequency (v_o) for the approaching and receding condensations) are given¹⁰ by

$$\delta_{r,a} = \frac{v_{r,a}}{v_o} = \gamma^{-1}(1 \pm \beta \cos \theta)^{-1} \quad (5)$$

where $\gamma = (1 - \beta^2)^{-1/2}$ is the Lorentz factor. As we know $\beta \cos \theta$ from equation (3), a determination of either v_a/v_o or v_r/v_o will allow the determination of β from equation (5) and thus the determination of θ and of the distance from equation (4). The spectral lines arising in the approaching and receding components should have mean redshifts of 0.75 and 2.36 respectively (as a result of relativistic effects, both should appear redshifted). Because of strong interstellar absorption, the search for line emission has to be done either in the infrared, where a variable counterpart with magnitude $K(2.2 \mu\text{m}) = 13\text{--}14$ was detected⁴, or in the hard X-rays.

The observations of this superluminal microquasar, together with the synchrotron radio-jets we found associated with the two persistent hard X-ray sources in the Galactic Centre region (1E1740.7–2942 and GRS1758–258, refs 17, 18 and 19, respectively) suggest that, irrespective of the masses of the collapsed objects, sources of hard X-rays and γ -rays usually eject clouds of plasma at relativistic velocities. Therefore one can expect to find ejecta with superluminal expansions not only in distant quasars, but also in high-energy sources of stellar mass. $\square$

Received 21 June; accepted 10 August 1994.

1. Porcas, R. W. in *Superluminal Radio Sources* (eds Zensus, J. A. & Pearson, T. J.) 12–25 (Cambridge Univ. Press, 1987).
2. Rees, M. J. *Nature* **211**, 468–470 (1966).
3. Kellermann, K. I. & Owen, F. N. in *Galactic and Extragalactic Radio Astronomy* (eds Verschuur, G. L. & Kellermann, K. I.) 563–602 (Springer, New York, 1988).
4. Mirabel, I. F. et al. *Astr. Astrophys.* **282**, L17–L20 (1984).
5. Harmon, A. et al. in *AIP Conf. Proc.* No. 304 (eds Fichtel, C. E., Gehrels, N. & Norris, J. P.) 210–219 (Am. Inst. of Physics, New York, 1994).
6. Castro-Tirado, A. et al. *Astrophys. J. Suppl. Ser.* **92**, 469–472 (1994).
7. Rodríguez, L. F. & Mirabel, I. F. *IAU Circ. No.* 5900 (1993).
8. Gerard, E., Rodríguez, L. F. & Mirabel, I. F. *IAU Circ. No.* 5958 (1994).
9. Sazonov, S., Sunyaev, R. & Lapshov, I. *IAU Circ. No.* 5959 (1994).
10. Pearson, T. J. & Zensus, J. A. in *Superluminal Radio Sources* (eds Zensus, J. A. & Pearson, T. J.) 1–11 (Cambridge Univ. Press, 1987).
11. Downes, D., Wilson, T. L., Bieging, J. & Wink, J. *Astr. Astrophys. Suppl.* **40**, 379–394 (1980).

12. Kouveliotou, C. et al. *Nature* **362**, 728–730 (1993).
13. Grindlay, J. E. *Astrophys. J. Suppl. Ser.* **92**, 465–468 (1994).
14. Hurley, K. et al. *Astrophys. J.* (in the press).
15. Finoguenov, F. et al. *Astrophys. J.* **424**, 940–942 (1994).
16. Margon, B. A. *Rev. Astr. Astrophys.* **22**, 507–536 (1988).
17. Mirabel, I. F., Rodríguez, L. F., Cordier, B., Paul, J. & Lebrun, F. *Nature* **358**, 215–217 (1992).
18. Mirabel, I. F. & Rodríguez, L. F. in *AIP Conf. Proc.* No. 304 (eds Fichtel, C. E., Gehrels, N. & Norris, J. P.) 413–420 (Am. Inst. of Physics, New York, 1994).
19. Rodríguez, L. F., Mirabel, I. F. & Martí, J. *Astrophys. J.* **401**, L15–L18 (1994).

ACKNOWLEDGEMENTS. We appreciate the open policy of the National Radio Astronomy Observatory, operated by Associated Universities Inc. under cooperative agreement with the US NSF. We thank E. Gerard, W. M. Goss, R. M. Hjellming, K. Kellermann and A. Zensus for enlightening exchanges and discussions on the radio properties of this source, and A. Harmon and J. Paul for information on its X-ray properties. We also thank R. N. Henriksen and A. Konigl for comments on the theoretical aspects of this work. L.F.R. acknowledges support from DGAPA, UNAM and CONACyT, Mexico.

Flat pearls from biofabrication of organized composites on inorganic substrates

Monika Fritz*†§, Angela M. Belcher*‡§, Manfred Radmacher†, Deron A. Walters†§, Paul K. Hansma†§, Galen D. Stucky†§, Daniel E. Morse*§¶ & Stephen Mann*‡§¶

* Marine Biotechnology Center, Marine Science Institute,

† Department of Physics, ‡ Department of Chemistry and

§ Materials Research Laboratory, University of California,

Santa Barbara, California 93106, USA

¶ School of Chemistry, University of Bath, Bath BA2 7AY, UK

THE study of biomineralization is inspiring new approaches to the controlled fabrication of synthetic materials such as nanoparticles, polymer–mineral composites and templated crystals^{1–3}. Although this biomimetic approach is gaining momentum, the biological mechanisms involved in biomineralization remain relatively unexplored. One major reason for this is the difficulty of analysing biomineralization processes in their native dynamic state. Here we demonstrate that a highly organized composite material—a ‘flat pearl’—can be biofabricated on disks of glass, mica and MoS₂ inserted between the mantle and shell of *Haliotis rufescens* (red abalone). We show that the construction of this material is spatially and temporally regulated and proceeds through a developmental sequence that closely resembles that at the growth front of the natural shell. Recognition of the implanted inorganic surfaces by mantle cells apparently governs a switch, perhaps genetically controlled, from aragonite to calcite biomineralization. Once a partially oriented calcite–protein primer layer has been deposited, there is a switch back to the nucleation and assembly of columnar stacks of highly ordered aragonitic nacre. Thus the presence of an inorganic surface between the mantle and shell of the organism triggers a change in the nature of the mineral phase deposited.

Haliotis rufescens is a marine gastropod that produces an adult shell consisting of an outer red layer of prismatic calcite and an inner layer of aragonitic nacre⁴. The growth front is regulated by epithelial cells which secrete organic macromolecules and inorganic ions into the extrapallial space between the mantle and shell⁵. A central tenet of shell biomineralization is that the organic constituents—including highly acidic proteins—play an important role in the spatial and chemical control of crystal nucleation and growth, and microstructural fabrication⁶. Several calcium-binding polyanionic peptides have been isolated from *H. rufescens* at various stages of development⁷, but the location of these proteins in the adult shell was not known. Recent atomic force microscopy studies on the structure, growth and dissolution of abalone nacre⁸ suggest that mineral bridges between adjacent layers of the lamellar microstructure may be responsible for the long-range crystallographic coherence observed in this biomineral.

We show here that mineralization to form an organized composite microstructure is induced on flat disks of abiotic substrate implanted between the mantle (secretory epithelium) and the shell, and that this system provides unique access to the sequence of events governing this biofabrication. Although the pioneering studies of Wada, Watabe and their colleagues on pearl formation in oysters and other bivalve molluscs suggested^{9–12} a similar depositional sequence to that reported here, their studies were hindered by the relatively slow growth of the pearls (~400 nacreous layers per year) and the anatomical complexity of the capsular ‘pearl sac’ containing the secretory cells¹³. Furthermore, the microlaminae of pearls grown in bivalves (like those of the

related bivalve nacre) do not contain the conically tapered stacks of crystals uniquely characteristic of the growing surface of nacre produced by gastropod molluscs¹⁰. The first stages of deposition were observed on glass coverslips and other abiotic substrates inserted between the mantle and shell of bivalves, although the formation of organized microstructures was not reported^{14,15}. Our studies were performed with disks (~18 mm diameter) of glass (coverslips), mica and MoS₂ inserted carefully between the mantle and shell of living adult red abalones. Coverslips extracted after 14 d were extensively mineralized across the surface facing the mantle tissue (Fig. 1). Three distinct regions—a red outer band, an iridescent central zone and a green inner area—could be delineated. The size of these regions depended on the precise location of the coverslip, but in general the outer (distal) band was the smallest zone, being 1–3 mm in width. Visibly, the colour was characteristic of red abalone shells. By comparison, the central iridescent region was not pigmented but extensively mineralized, often being 6–10 mm in width and 300 µm in thickness after 14 d of insertion. The green inner (proximal) region was 8–10 mm in width and only partially mineralized, being primarily composed of a water-insoluble organic material.

X-ray diffraction measurements were undertaken to characterize the mineral composition of each of the three zones. The results indicated that whereas both the red and green regions were mineralized with calcite, the iridescent zone contained aragonite oriented along the (001) crystallographic plane (index of orientation $r=0.065$ based on Rietveld refinement¹⁶) (Fig. 2a). Scanning electron micrographs of the iridescent zone showed a laminated microstructure of pseudo-hexagonal aragonite tablets (10–20 µm wide × 0.5 µm thick) intercalated between organic polymeric sheets (Fig. 3a). Over 200 layers of these organized crystals were observed in samples removed after 14 d. Because nacre deposition began only after day ~7, this represents a growth rate of ~200 nacreous layers per 7 d; that is, ~26 times faster than the average rate of nacreous growth in pearls in

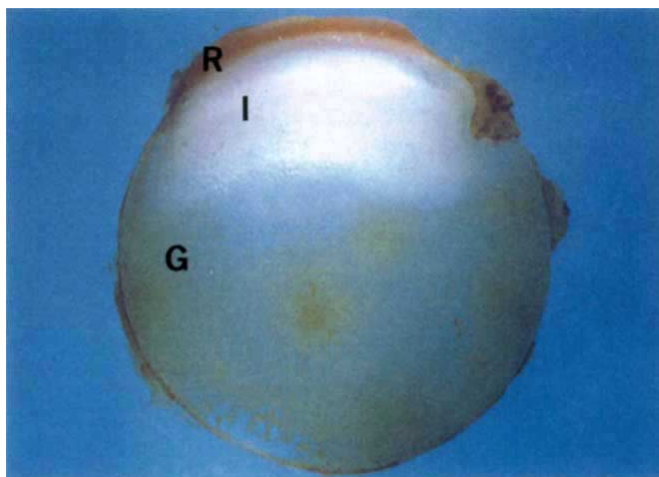
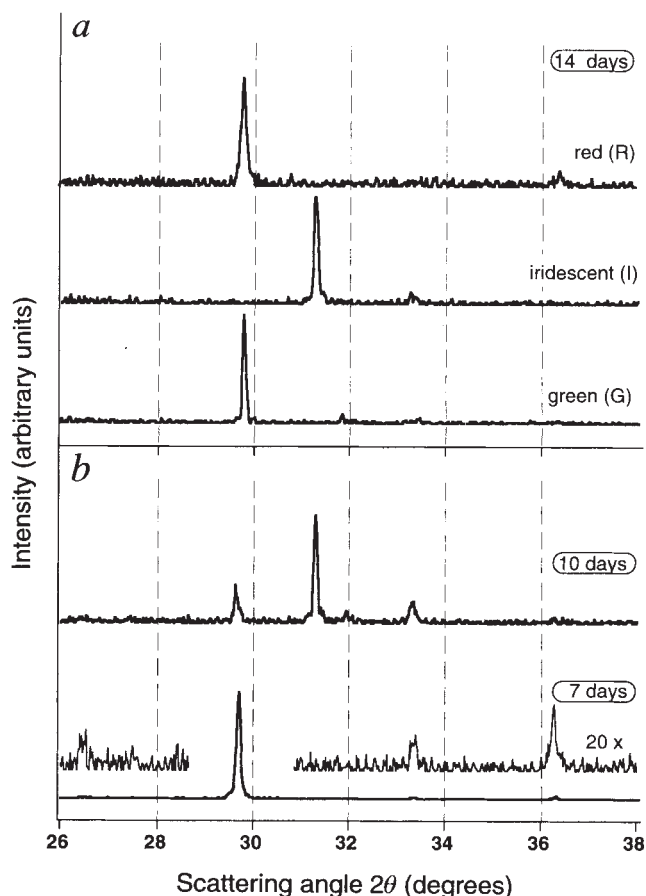


FIG. 1 Optical micrograph showing a ‘flat pearl’ deposited on an 18 mm diameter glass coverslip inserted between the mantle and shell of a 4-year-old adult red abalone (~120 mm long). Three distinct regions, mirroring those in the natural shell, can be identified. The distal red zone (R) was formed by epithelial cells at the mantle edge. The iridescent (I) middle region resembles the mother-of-pearl inner layer (nacre) of the natural shell and the green area (G) consists primarily of pigmented organic material.

METHODS. Animals were kept at 15 °C in fresh flowing sea water in a 150-l tank (replacement rate: 50% per hour) and fed on the kelp *Macrocystis pyrifera*. After manual insertion of the coverslips, animals were kept undisturbed in the tank for up to 14 d. The coverslips were carefully removed by hand and kept in ultraviolet-sterilized sea water before SEM and X-ray diffraction analysis.

¶ To whom correspondence should be addressed.



bivalve molluscs. Images of the growth front showed discrete columnar stacks of aragonite crystals, with the most recently deposited crystals exhibiting progressively smaller lateral dimensions (Fig. 3a)—a constructional pattern characteristic of gastropod nacre⁴. Significantly, the interface between the glass coverslip and this nacreous material in the iridescent region did not consist of oriented nacre, but was found to be a calcified 100- μ m layer of prismatic material similar to that observed in the red outer region of the coverslip. X-ray diffraction studies of a 14-d-old sample mounted with this underlying mineral toward the incident beam (that is, shell-side up) indicated that this interface comprised a basal layer of calcite.

Elucidation of the temporal sequence of microstructural development of the iridescent zone was undertaken by removing mineralized coverslips after different implantation times. The first mineral to be deposited was calcite, which after 5 d was organized into 100- μ m intergrown crystalline blocks. These calcite crystals showed orientation about the $\{104\}$ crystallographic plane (index of orientation $r=0.3$ based on Rietveld refinement¹⁶), with the other planes oriented randomly (Fig. 3b). Aragonite nucleation then occurred after 7 d at discrete sites on the surface of this calcite primer or its associated organic molecules (Fig. 3c). Although these crystals appeared not to be highly oriented (Fig. 2b), they were progenitors of discrete columnar stacks of oriented aragonite tablets which subsequently became interdigitated by lateral growth during development of the nacreous architecture.

The results indicate that biomineralization on abiotic substrates can be spatially and temporally organized such that the patterning of the composite closely resembles the growth front of the natural shell. Moreover, biofabrication of the 'flat pearl' involves the secretion of specific proteins from the adjacent mantle epithelial cells. We have isolated at least nine distinct water-soluble proteins from the demineralized composite, corresponding to proteins previously characterized from the natural

FIG. 2 X-ray diffraction data of mineralized overgrowths on glass coverslips. **a**, Spatial distribution of calcite and aragonite minerals in a 14-d calcified overgrowth. Data were recorded from material isolated from the areas R, I and G in Fig. 1. Reflections corresponding to the mineral calcite (2θ and (hkl) values; 23.09° (102), 29.414° (104), 36.005° (110), 39.483° (113), 43.23° (202), 47.535° (108), 48.554° (116)) were observed in measurements obtained from the red distal region. Thus, this region was similar both in colour and mineralogy to the prismatic outer layer of the natural shell. In contrast, the iridescent (I) area gave reflections consistent with the polymorph, aragonite (2θ and (hkl) values; 31.12° (002), 33.13° (012)), and the presence of only two reflections indicated that the crystals were preferentially oriented approximately along their c axes. The green proximal region was lightly mineralized with calcite (only the most intense (104) reflection at $2\theta = 29.41^\circ$ was observed). **b**, Temporal sequence of mineralization of the developing iridescent region on glass coverslips. After 7 d of insertion, X-ray reflections for both calcite (2θ and (hkl) values; 29.2° (104), 36.00° (110) and a minor phase of non-oriented aragonite (2θ and (hkl) values; 26.25° (111), 27.31° (021), 33.13° (012), 37.93° (112), 38.51° (130), 38.69° (022) were present. After 10 d, mainly oriented aragonite was observed (2θ and (hkl) values; 31.12° (002), 33.13° (012)), with some calcite (2θ and (hkl) values; 29.414° (104)). After 14 d only oriented aragonite was observed in samples analysed mantle-side up (see Fig. 2a).

METHODS. A 14-d calcified overgrowth was removed intact from the substrate by cutting and gently breaking the coverslip from the unmineralized side. Representative areas were removed with a diamond saw and the resulting samples mounted mantle-side up for X-ray diffraction analysis. Mineralized overgrowths at earlier stages of development were studied *in situ* on intact coverslips. X-ray diffraction measurements were performed with the mantle-side up (facing the incident beam) using a Scintag PADX diffractometer equipped with a liquid-nitrogen-cooled germanium solid-state detector using Cu K α radiation.

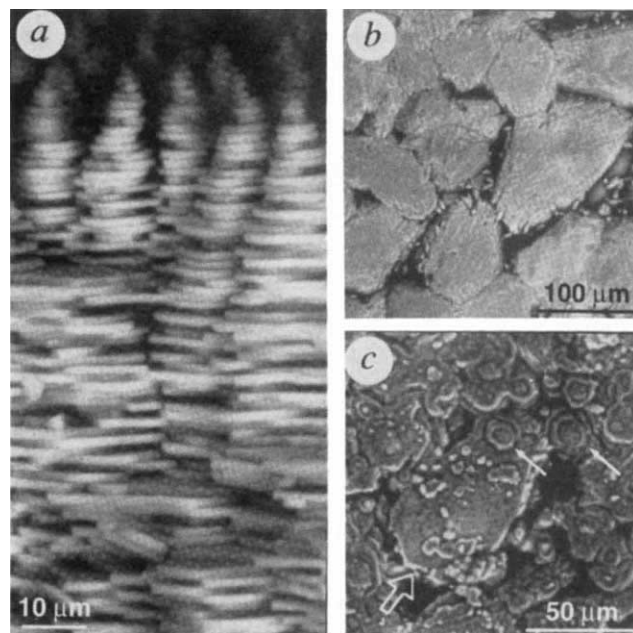
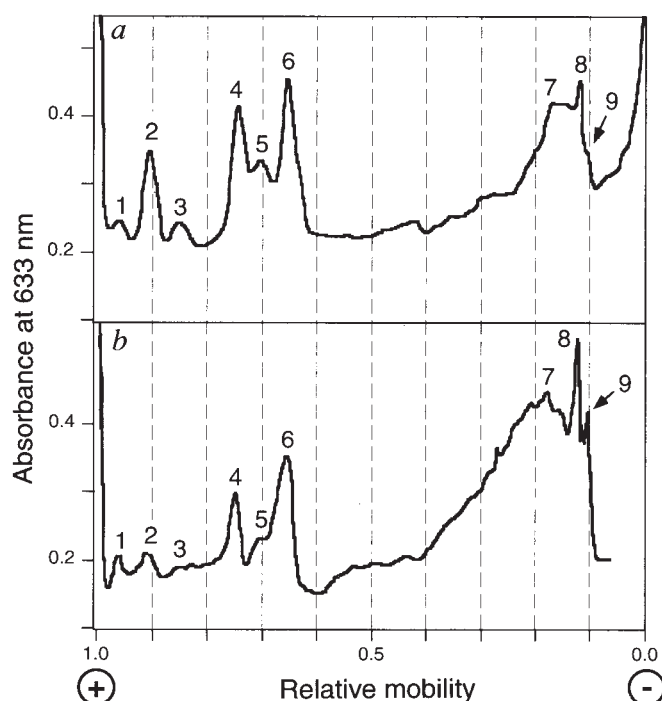


FIG. 3 Scanning electron microscopy (SEM) images showing the ultrastructure and development of 'flat pearls'. **a**, Image of the mature deposit (14 d after insertion) showing the laminated microstructure and columnar growth front. **b**, Initial deposit of calcite formed at the substrate interface after 5 d. **c**, Nucleation of aragonite crystals (arrows) on the calcitic basal layer (open arrow) after 7 d.

METHODS. Samples were blown-dry immediately before being mounted on a 45° SEM stage without further preparation. Images were collected in a Jeol JSM5300LV environmental SEM in the low-vacuum mode at 15 kV with a detector-sample separation of 15 mm. The pressure was 200 mtorr and the images were collected in the backscattered electron mode.



shell (Fig. 4). Because these proteins are all highly acidic, optimal detection after electrophoretic resolution is achieved by staining with a cationic carbocyanin dye⁷. Although gels stained with this dye exhibit poor photographic contrast, the positions of the stained proteins are reproducible and readily detected by laser densitometry¹⁷. It can be seen that the electrophoretic mobilities of these highly acidic proteins match closely those of the proteins extracted from the intact shell, although the relative staining intensities of the individual components differ in the two samples. (Thus, for example, peaks no. 3 and no. 5 are pronounced in Fig. 4a but seen only as shoulders in Fig. 4b; peak no. 9 is pronounced in Fig. 4b, but only a shoulder in Fig. 4a.) These differences in relative intensity reflect the fact that some of the proteins are unique to the calcitic composites, whereas others are unique constituents of the aragonitic nacre (A.M.B. and D.E.M., unpublished observations), and the fact that the samples analysed contain different proportions of these two mineralized composites (as verified by X-ray diffraction data). None of the species detected in Fig. 4 corresponds to a molecule that is not a protein. Electrophoresis under denaturing conditions reveals that the proteins extracted from the 'flat pearls' have relative molecular masses indistinguishable from those of the proteins isolated from the intact shell (A.M.B. and D.E.M., unpublished data).

Exposure of the mantle epithelial cells to foreign glass or single-crystal substrates, including mica {001} and MoS₂ {0001}, stimulates a cellular response that initially activates a similar sequence of protein secretion and calcite mineralization, followed by nacre deposition. The findings that (1) the same depositional sequence is observed on glass, mica and molybdenum sulphide surfaces, which exhibit a wide range of surface properties, and (2) this depositional sequence is identical to that observed in the native shell, exclude the possibility that the sequence we observed was induced as a unique response to the hydrophilic glass coverslip surface. We have observed no evidence of withdrawal of the mantle edge, suggesting that epithelial cells apposed to areas of nacre deposition in the natural shell essentially remain spatially fixed in the presence of the coverslips. Thus, we propose that these cells respond by down-regulation of nacre production until the calcite-protein matrix is established on the abiotic surface. These observations are consistent with

FIG. 4 Gel electrophoresis of protein extracted from 'flat pearl' (a) and natural shell (b).

METHODS. These are modifications of those described previously¹⁷. Samples of 'flat pearl' and shell were ground with mortar and pestle; soluble proteins were extracted from 0.1 g of the resulting powders by demineralization with ethylenediaminetetraacetic acid (EDTA, 0.5 M, in 0.01% w/v NaNO₃, adjusted to pH 8.0 with NaOH; final volume, 1.5 ml) agitated for 12 h at room temperature. The solutions then were adjusted to 5 ml and 0.4 M NH₄HCO₃, and the proteins concentrated by centrifugal dialysis (Centriprep concentrator). The soluble proteins were resolved by electrophoresis on a 20% polyacrylamide gel (1 mm thick, 8 cm long) under non-denaturing conditions, stained with 1-ethyl-2-[3-(1-ethylnaphthol [1,2-d]-thiazolin-2-ylidene)-2-methylpropenyl]-naphtho [1,2-d] thiazolium bromide and the gel scanned with an LKB Ultrascan XL laser densitometer.

information obtained on the process of shell regeneration in molluscs subjected to damage^{18,20}. Under these conditions, epithelial cells engaged in nacre deposition adopt processes normally restricted to cells at the mantle edge. We postulate, therefore, that the biological fidelity of molluscan shell architecture is a consequence of a dynamic interaction at the cell-mineral interface, in which cell recognition of the abiotic surface governs a genetic switch¹⁷ controlling the structure of the mineral that is formed. Although the nature of this interfacial recognition is not known, it is clear that severing the communication between mantle and mineral, for example by insertion of a coverslip in *H. rufescens*, results in a defaulting of cell activity and a re-initiation of the mineralization sequence (that is, deposition of an oriented calcite primer, followed by aragonitic nacre). Further studies of 'flat pearl' formation in gastropods should provide additional insights into the mechanisms governing the biosynthesis of organized composite materials. □

Received 8 April; accepted 2 August 1994.

1. Mann, S. *Nature* **365**, 499–505 (1993).
2. Mann, S. et al. *Science* **261**, 1286–1292 (1993).
3. Heuer, A. H. *Science* **255**, 1098–1105 (1992).
4. Nakahara, H., Bevilander, G. & Kakei, M. *Venus* **41**, 33–46 (1982).
5. Simkiss, K. & Wilbur, K. M. in *Biomineralization* (eds Simkiss, K. & Wilbur, K. M.) (Academic, San Diego, 1989).
6. Addadi, L. & Weiner, S. *Angew. Chem. int. Edn engl.* **31**, 153–161 (1992).
7. Cariolou, M. A. & Morse, D. E. *J. comp. Physiol.* **B157**, 717–729 (1988).
8. Manne, S. et al. *Proc. R. Soc. B* **256**, 17–23 (1994).
9. Watabe, N. in *Biology of the Integument 1* (eds Bereiter-Hahn, J., Matoltz, A. G. & Richards, K. S.) 448–485 (Springer, New York, 1984).
10. Wada, K. *Bull. Natn. Pearl Res. Lab.* **7**, 703–827 (1961).
11. Simkiss, K. & Wada, K. *Endeavour* **4**, 32–37 (1980).
12. Götting, K. J. *Experientia* **35**, 756 (1979).
13. Wada, K. *Bull. Jap. Soc. Sci. Fish.* **23**, 302 (1957).
14. Watabe, N. & Wilbur, K. M. *Nature* **198**, 334 (1960).
15. Wada, K. *Bull. Jap. Soc. Sci. Fish.* **26**, 549–553 (1960).
16. Rietveld, H. M. J. *appl. Crystallogr.* **2**, 65–71 (1969).
17. Morse, D. E., Cariolou, M. A., Stucky, G. D., Zaremba, C. M. & Hansma, P. K. *Mater. Res. Soc. Symp. Proc.* **292**, 59–67 (1993).
18. Meenakshi, V. R., Blackwelder, P. L. & Wilbur, K. M. *J. Zool.* **171**, 457–484 (1973).
19. Meenakshi, V. R., Martin, A. W. & Wilbur, K. M. *Mar. Biol.* **27**, 27–35 (1974).
20. Blackwelder, P. L. & Watabe, N. *Biomineralization* **9**, 1–10 (1977).

ACKNOWLEDGEMENTS. We thank C. Zaremba, J. Cleveland and R. Giles for discussions. This work was supported by grants from the US Office of Naval Research (D.E.M., G.D.S. and P.K.H.), the National Science Foundation, through the Materials Research Laboratory at the University of California, Santa Barbara (A.M.B., M.R., D.A.W., G.D.S., D.E.M. and P.K.H.), and a fellowship from the Deutsche Forschungsgemeinschaft (M.F.).

Feedbacks between climate and boreal forests during the Holocene epoch

Jonathan A. Foley, John E. Kutzbach,
Michael T. Coe & Samuel Levis

Institute for Environmental Studies, and Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, 1225 West Dayton Street, Madison, Wisconsin 53706, USA

PREVIOUS studies^{1–5} have demonstrated that the predictions of global climate models are highly sensitive to large changes in vegetation cover, such as the complete removal of tropical or boreal forests. Although these studies have illustrated the potential effects of massive deforestation on the climate system, vegetation changes of this scale are very unlikely to occur. Investigating past environments may better illustrate the possible interactions between climate and vegetation cover. For example, palaeobotanical evidence indicates that 6,000 years ago boreal forests extended north of the modern treeline⁶, apparently in response to high-latitude warming resulting from variations in the Earth's orbit^{7,8}. The expanded boreal forests, which took the place of tundra, must also have affected climate by significantly reducing the surface albedo⁵. Here we use a global climate model to examine the relative effects of orbitally-induced insolation variations and of the northward extension of boreal forests on the mid-Holocene climate. Orbital variations alone warm the high latitudes by 2 °C or more in summer, autumn and winter. The subsequent northward extension of boreal forests gives rise to an additional warming of approximately 4 °C in spring and about 1 °C in the other seasons. This suggests that large positive feedbacks between climate and boreal forests may have taken place in the recent geological past.

We perform three climate simulations with the GENESIS global climate model (GCM) version 1.02 (refs. 9, 10): Control (control climate simulation), Exp-A (climate simulation with changes in the Earth's orbit) and Exp-B (climate simulation with changes in the Earth's orbit and expanded boreal forests). In all of these simulations atmospheric CO₂ concentrations are held constant. We examine the effects of two forcing mechanisms on the global climate: orbital forcing (Exp-A minus Control) and the subsequent vegetation forcing (Exp-B minus Exp-A).

In Exp-A, we change the Earth's orbital parameters from the modern values to represent conditions 6 kyr ago; eccentricity from 0.0167 to 0.0187, axial tilt from 23.4° to 24.1° and date of perihelion from early January to mid-September^{7,8}. The change in the date of perihelion increases the amplitude of the seasonal cycle of Northern Hemisphere insolation by ~5%, leading to enhanced heating in summer and enhanced cooling in winter. The increased axial tilt further enhances high-latitude summer insolation. In the high northern latitudes, increased summer insolation leads to warmer ocean temperatures and delays the formation of sea ice, resulting in warmer winter conditions and hence year-round warming¹¹.

Palaeobotanical data indicate that a northward expansion of boreal forests occurred during the middle Holocene^{12–19}. This expansion must have resulted, in part, from the orbitally-induced high-latitude warming¹⁴. In Exp-B, we extend the northern limit of boreal forests based on two criteria: (1) palaeobotanical evidence, which is sparse in many regions, and (2) GCM sensitivity considerations which require that changes in boundary conditions cover several contiguous model grid-cells (Fig. 1). Because of the coarse spatial resolution of the GCM (~4.5° latitude by 7.5° longitude), there are differences between the imposed boreal forest expansion and the palaeobotanical data. For example, in portions of North America, the observed expansion was less extensive than that imposed in our simulations; but in Asia, there is better agreement. In the future it will be possible to repeat these simulations with higher-resolution GCMs and more complete palaeovegetation datasets.

Both orbital forcing and vegetation forcing operate mainly by affecting the amount of solar radiation absorbed by the surface. In response to orbital forcing, annual absorbed solar radiation in the high northern latitudes (defined here as 60–90° N) increases by ~4 W m⁻² (Table 1). On the annual average, vegetation forcing lowers the land surface albedo from 0.36 to 0.26 (Table 1), primarily by replacing snow-covered tundra with an evergreen forest that protrudes above the snow, resulting in an additional 8 W m⁻² absorbed solar radiation. In these simulations, the changes in high-latitude cloud cover are not statistically significant and do not appreciably feed back on the surface energy balance.

On the annual average, orbital forcing warms the high latitudes by 1.8 °C (Table 1, Fig. 2a). Vegetation forcing produces an additional annual-average warming of 1.6 °C (Table 1, Fig. 2b). The largest temperature increases due to vegetation forcing generally coincide with the areas of expanded forest.

FIG. 1 The prescribed extent of boreal forests used in the GENESIS GCM simulations. GENESIS specifies surface properties on a 2° latitude by 2° longitude grid (solid colours), whereas the atmospheric model operates on a 4.5° latitude by 7.5° longitude grid (lines).

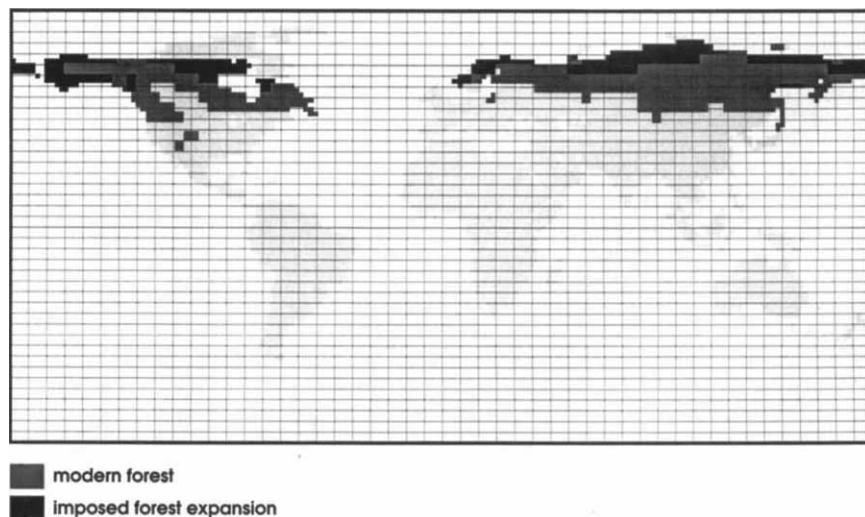


TABLE 1 Simulated climate variables from 60 to 90 °N

	Control	Exp-A	Exp-B
Surface temperature (°C)	-8.4	-6.6	-5.0
Deep soil ice fraction (1.75–4.25 m)	0.31	0.28	0.23
Precipitation (mm d ⁻¹)	1.73	1.81	1.88
Runoff (mm d ⁻¹)	0.89	1.00	1.09
Sea ice thickness (cm)	79.0	63.0	56.0
Sea ice fraction	0.52	0.46	0.43
Snow-cover thickness (cm H ₂ O)	13.0	11.0	10.0
Snow-cover fraction	0.40	0.34	0.30
Surface albedo, winter	0.434	0.423	0.301
Surface albedo, spring	0.576	0.581	0.414
Surface albedo, summer	0.172	0.166	0.149
Surface albedo, autumn	0.275	0.234	0.182
Surface albedo, annual	0.364	0.356	0.265
Absorbed solar radiation, winter (W m ⁻²)	7.3	6.0	6.9
Absorbed solar radiation, spring (W m ⁻²)	136.0	135.1	157.1
Absorbed solar radiation, summer (W m ⁻²)	253.9	277.4	284.8
Absorbed solar radiation, autumn (W m ⁻²)	43.8	49.6	50.9
Absorbed solar radiation, annual (W m ⁻²)	111.2	115.3	123.3

Summary of GENESIS climate model simulations for Control (modern), Exp-A (orbital forcing) and Exp-B (orbital forcing + vegetation forcing). All values are annual averages for land only unless otherwise stated. As a general basis of comparison of the sensitivity of this climate model, we note that the GENESIS GCM simulates a global surface temperature warming of 2.1 °C due to doubled atmospheric CO₂, which is towards the low end of the range in recent climate model experiments²³.

However, the magnitude and pattern of these changes must be assessed relative to the inherent interannual variability of the GCM²⁰. The simulated interannual variability of surface temperature is large across North America from eastern Alaska to Greenland, and small across northern Eurasia. Thus the simulated warming due to vegetation forcing (Fig. 2b) has a higher statistical significance over northern Eurasia than over eastern North America and Greenland. Similarly, the 3 °C warming due to orbital forcing (Fig. 2a) in North America has no higher statistical significance than the 1–2 °C warming elsewhere. Overall, the 1.8 °C annual-average warming due to orbital forcing, and the additional 1.6 °C warming due to vegetation forcing, are both statistically significant at the 95% level.

Although the annual-average temperature changes from both orbital and vegetation forcing are of similar magnitude, the seasonal responses differ significantly (Fig. 3). Under orbital forcing, monthly average temperature changes are >1 °C beginning in June, with a maximum change of ~5 °C in December. Vegetation forcing produces a maximum temperature response of ~4 °C in March and April, and temperature increases of >1 °C occur throughout the year (except in January and October).

Snow and ice cover play an important role in the high-latitude climates. Both orbital forcing and the subsequent vegetation forcing result in similar decreases in snow cover over land and Arctic sea ice (Table 1). Taken together, orbital and vegetation forcing reduce snow and sea-ice volume by nearly 40%. The reduction of the highly reflective snow and ice cover enhances both orbital and vegetation forcing, thus acting as an additional positive feedback mechanism⁵.

The hydrological cycle is strongly affected by both orbital and vegetation forcing. Annual precipitation over high-latitude land increases by ~5% in both simulations (Table 1). The enhanced rate of precipitation results from an increase in the water vapour content of the atmosphere under warmer conditions. Runoff,

which includes surface runoff and subsurface drainage, increases by ~10% in response to both orbital and vegetation forcing. In addition, orbital and vegetation forcing each warm the high-latitude soils and reduce the amount of deep soil ice by ~10%.

Compared to the present day, the climate in the middle Holocene was substantially warmer in the high northern latitudes. Our mid-Holocene climate simulation without vegetation feedbacks (Exp-A) yields a 1.8 °C high-latitude warming, in good agreement with other modelling studies^{7,8,11}. However, palaeoclimatic analyses indicate that the annual-average surface temperatures may have been as much as 3 °C warmer than present in northern Eurasia and northern North America in the middle Holocene^{21,22}. The large positive feedback due to the northward extension of boreal forests (Exp-B) could help to explain the amount of observed mid-Holocene warmth, which orbital forcing alone does not appear to produce.

The amplification of high-latitude warming shown here may also be relevant to issues of anthropogenic climate change. The

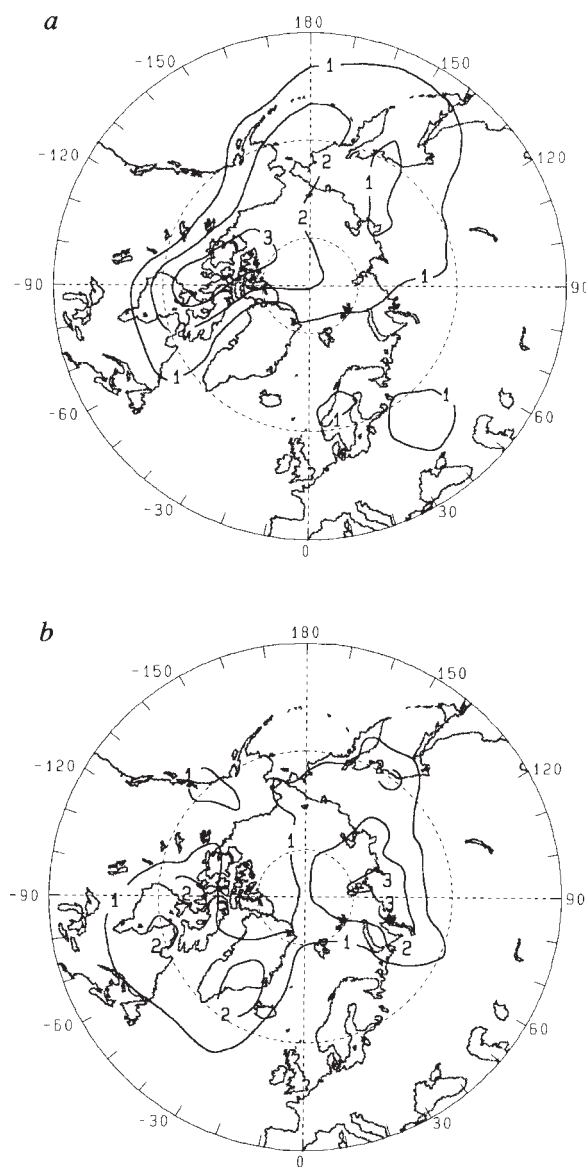


FIG. 2 a, Annual-average surface temperature increases due to mid-Holocene orbital forcing (Exp-A minus Control). Units are in °C; the contour interval is 1 °C. b, Annual-average surface temperature increases due to mid-Holocene vegetation forcing (Exp-B minus Exp-A). Same units and contour interval as a.

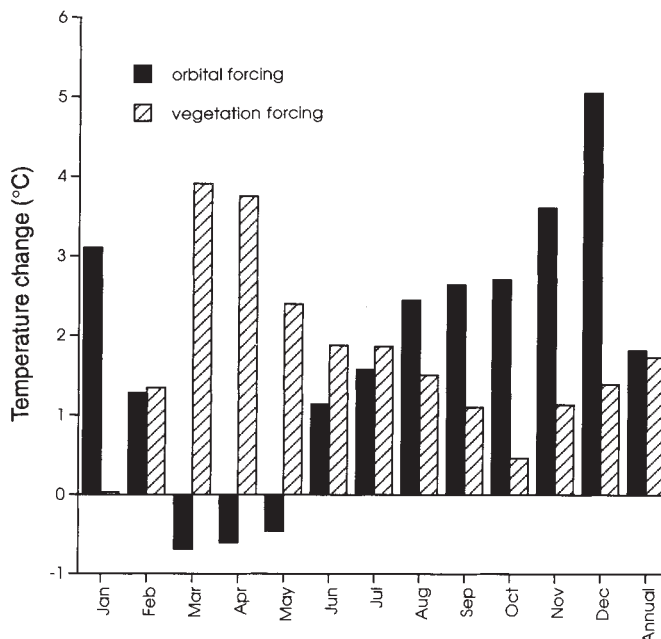


FIG. 3 Changes in zonal-average (over land from 60° to 90° N) monthly surface temperatures due to orbital and vegetation forcing.

estimates of high-latitude warming in response to greenhouse gas emissions²³ may need to be re-evaluated for the effects of possible vegetation feedbacks in light of these results. Further investigations of the interactions between climate and vegetation should be performed with fully coupled models of the climate system (including improved cloud, sea-ice and ocean processes) and the terrestrial biosphere. In addition, improvements in model resolution would depict the shifts in palaeovegetation more accurately than was possible here. □

A cellular model of braided rivers

A. Brad Murray & Chris Paola

Department of Geology and Geophysics, University of Minnesota, Minneapolis, Minnesota 55455, USA

A BROAD sheet of water flowing over non-cohesive sediment typically breaks up into a network of interconnected channels called a braided stream (Fig. 1). The dynamics of such networks are complex; channels migrate laterally, split, rejoin and develop bars, with the flow shifting unpredictably from one part of the network to another. Many processes are known to operate in a braided river¹⁻³, but it is unclear which of these are essential to explain the observed dynamics. We describe here a simple, deterministic numerical model of water flow over a cohesionless bed that captures the main spatial and temporal features of real braided rivers. The patterns arise from local scour and deposition caused by a nonlinear dependence of bedload sediment flux on water discharge. Although the morphology of the resulting network depends in detail on the sediment-transport rule used in the model, our results suggest that the only factors essential for braiding are bedload sediment transport and laterally unconstrained free-surface flow.

Recently, a number of workers have used cellular computer models to study naturally occurring spatial patterns⁴⁻⁹. In these models the cells of a lattice interact according to rules based on abstractions of the physics governing a system. Here we present a cellular model of stream braiding (see Box 1). The initial condition for the model is a uniform slope with white-noise elevation perturbations; the latter have an amplitude of the order of the mean downhill elevation difference between adjacent rows of cells. We include high side walls to contain the flow. An iteration begins when water is introduced into some of the cells at the upstream end of the lattice, and begins moving downstream cell-by-cell, carrying sediment. When the water reaches the downstream end of the lattice, the iteration ends, and the elevation

Received 16 March; accepted 28 July 1994.

- Dickinson, R. E. & Henderson-Sellers, A. Q. *Jl R. met. Soc.* **114**, 439-462 (1988).
- Shukla, J., Nobre, C. & Sellers, P. J. *Science* **247**, 1322-1325 (1990).
- Henderson-Sellers, A. et al. *J. geophys. Res.* **98**, 7289-7315 (1993).
- Lean, J. & Rowntree, P. R. Q. *Jl R. met. Soc.* **119**, 509-530 (1993).
- Bonan, G. B., Pollard, D. & Thompson, S. L. *Nature* **359**, 716-718 (1992).
- Ritchie, J. C. *Postglacial Vegetation of Canada* (Cambridge Univ. Press, 1987).
- COHMAP Members *Science* **241**, 1043-1052 (1988).
- Wright, H. E. Jr, et al. (eds) *Global Climates since the Last Glacial Maximum* (Univ. Minnesota Press, Minneapolis, 1993).
- Thompson, S. L. & Pollard, D. *J. Clim.* (in the press).
- Pollard, D. & Thompson, S. L. *Glob. planet. Change* (in the press).
- Mitchell, J. F. B., Grahame, N. S. & Needham, K. H. *J. geophys. Res.* **93**, 8283-8303 (1988).
- Bryson, R. A., Irving, W. N. & Larsen, J. A. *Science* **147**, 46-48 (1965).
- Sorenson, C. J., Knox, J. C., Larsen, J. A. & Bryson, R. A. *Quat. Res.* **19**, 1-17 (1972).
- Ritchie, J. C., Cwynar, L. C. & Spear, R. W. *Nature* **305**, 126-128 (1983).
- Khotinsky, N. A. in *Late Quaternary Environments of the Soviet Union* (eds Velichko, A. A., Wright, H. E. Jr & Barnosky, C. W. 179-200 (Univ. Minnesota Press, Minneapolis, 1984).
- Cwynar, L. C. & Spear, R. W. *Ecology* **72**, 202-212 (1991).
- Keenan, T. J. & Cwynar, L. C. *Can. J. Bot.* **70**, 1336-1345 (1992).
- Peterson, G. in *Global Climates Since the Last Glacial Maximum* (eds Wright, H. E. Jr et al.) (Univ. Minnesota Press, Minneapolis, 1994).
- Webb, T., Bartlein, P. J. & Kutzbach, J. E. in *North America and Adjacent Oceans During the Last Glaciation* (eds Ruddiman, W. F. & Wright, H. E. Jr) (Geological Soc. Am., Boulder, Colorado, 1987).
- Kutzbach, J. E. & Guetter, P. J. *J. atmos. Sci.* **43**, 1726-1759 (1986).
- Frenzel, B., Pécsi, M. & Velichko, A. A. (eds) *Atlas of Palaeoclimates and Palaeoenvironments of the Northern Hemisphere: Late Pleistocene-Holocene* (Geographical Res. Inst., Hungarian Acad. Sci. Budapest, 1992).
- Guiot, J., Harrison, S. P. & Prentice, I. C. *Quat. Res.* **40**, 139-149 (1993).
- Houghton, J. T., Jenkins, G. J. & Ephraums, J. J. (eds) *Climate Change: The IPCC Scientific Assessment* (Cambridge Univ. Press, 1990).

ACKNOWLEDGMENTS. We thank S. Thompson and D. Pollard for providing the GENESIS climate model. We also thank P. Behling, R. Gallimore and R. Selin for assistance with the simulations. This manuscript was substantially improved by the comments of J. Mitchell, I. C. Prentice and T. Webb. This research was supported by the US NSF the US DOE and the Climate, People and Environment Program (CPEP) of the University of Wisconsin. The computer simulations were performed at the National Center for Atmospheric Research, which is supported by the US NSF.



FIG. 1 Aerial photograph of part of the braided Aichilik River on the north slope of the Brooks Range in Alaska. Flow is from top to bottom. The slope is ~0.001, the gravel flat is ~0.5 km wide, and the flow depths are of the order of 1 m.

of each cell is adjusted according to the difference between the total sediment loads entering and leaving the cell (a discretization of the Exner mass-balance equation). The elevations of the first and last rows of cells remain unchanged, simulating the rigid end walls of laboratory experiments. There is no net gain or loss of sediment in the model runs.

From each cell, water is distributed to any (or all) of the three downstream immediate neighbours. (We route the water only to the downstream neighbours—rather than to any adjacent cell that has a lower elevation—because it simplifies the algorithm, and because we have rarely observed channels in braided streams that flow at an angle $>45^\circ$ to the overall downhill direction.) The amount going to each neighbour is proportional to S^n where S is the local (directional) slope and n varies between 0.5 and 1. The resulting patterns do not seem to be sensitive to the exact value of n . This simple water routing scheme does not explicitly include backwater effects, although the fact that information cannot propagate upstream in the flow implies supercritical flow everywhere.

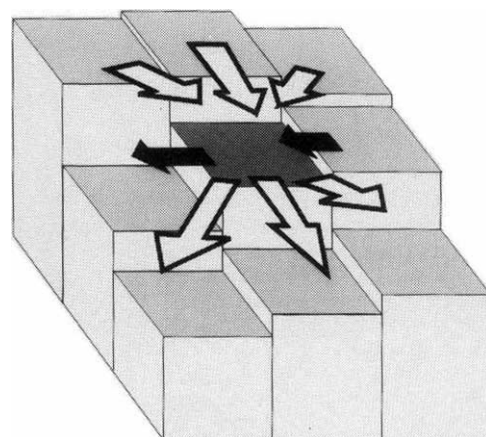
We have used several sediment-transport relationships. In the simplest, the sediment flux is proportional to Q^m where Q is the water discharge and $m > 1$ (rule 1). This leads to the basic instability when starting from a bed with random perturbations: because an isolated low area captures more of the flow than the surroundings, this nonlinear dependence on flow strength makes the sediment transport greater in the low area than in all the areas feeding into it combined. This creates a zone in which more sediment is leaving than entering (that is, the sediment flux diverges), and leads to a deepening of the low area. The converse occurs over an isolated high area, causing the flow to split around the growing bar. We, as well as other workers^{10,11}, have observed these instabilities operating in scale-model laboratory experiments. Starting from a nearly smooth surface, isolated scour holes form, and the scoured material is deposited nearby. These holes and deposits then divert the flow into each other, eventually coalescing into the channels and bars of a braided stream. The model develops in a similar fashion with most combinations of rules and parameters we have tried (we discuss exceptions below).

This simple discharge-dependent sediment transport produces a fully developed pattern that is crudely braided—channels split and rejoin—but unrealistic in that downstream width changes are extremely abrupt. We can better approximate empirical transport laws by letting sediment flux depend on stream-power index QS as $(QS)^m$ (rule 2). Based on Ashmore's¹² compilation of natural and laboratory gravel-bed rivers, a reasonable value for m is 2.5. The results do not depend sensitively on this value, as long as $m > 1$. When starting from a nearly smooth bed, this transport rule tends to dampen the initial perturbations, leading to a progressively smoother bed, unless $n \geq 1$ and the initial perturbations have an amplitude at least several times the average elevation difference between rows. These conditions also produce a crudely braided pattern that is qualitatively unrealistic in the ways described above. The above relationship (rule 2) between sediment flux and stream power is based on spatially averaged data. Applied locally, it gives the unrealistic result that sediment flux is zero where the local slope is zero. In real rivers, however, sediment transport commonly occurs where local bed slopes are zero or negative (uphill), as long as the water continues to flow downstream^{13,14}. To include this effect, we add a term to the stream power that is proportional only to discharge (rule 3). This formulation, with a constant of proportionality equal to 2 or 3 times the average slope, produces a realistic fully developed pattern (Fig. 2). We have also done simulations in which the uphill-transport term depends on the discharges and slopes immediately upstream (rule 4). This also leads to a realistic pattern.

In real rivers, when the slope is not parallel to the flow direction (for example, near channel banks), gravity produces a lateral downhill component in the sediment transport^{15,17}. We

BOX 1 Rules used in the cellular model

The arrangement of the cells is shown below. White arrows represent the discharge and sediment flux from cell to cell, with lengths indi-



cating magnitudes. If the slope to at least one of the three downstream neighbours is positive, we route the water only into that (those) cell(s), according to: $Q_i = Q_0 S_i^n / \sum S_i^n$ where Q_i is the discharge and S_i is the slope from the cell in question into downstream neighbour i (where $i = 1, 2$, and 3 corresponds to left, centre and right, respectively) and Q_0 is the total discharge in the cell we are routing water from. The sum, which runs over all the neighbours with positive slopes, normalizes the water routing, so that all the discharge in a cell is routed downstream. In its simplest form, the equation of motion for flow suggests that n should be 0.5, so we usually use that value, although we have also tried $n = 1$. Slopes equal the elevation difference to each neighbour, divided by $\sqrt{2}$ for the diagonal neighbours. If none of the slopes to the immediate downstream neighbours are positive, but at least one slope is 0, the water is distributed evenly to this (these) neighbour(s). If all the slopes are negative, then we distribute the water to all three according to: $Q_i = Q_0 |S_i|^{-n} / \sum |S_i|^{-n}$ (water can flow uphill in real rivers when either the surface slope is positive or the momentum of the water is great enough). We have used four sediment-transport rules:

$$Q_{si} = K Q_i^m \quad (1)$$

$$Q_{si} = K (Q_i S_i)^m \quad (2)$$

$$Q_{si} = K [Q_i (S_i + C)]^m \quad (3)$$

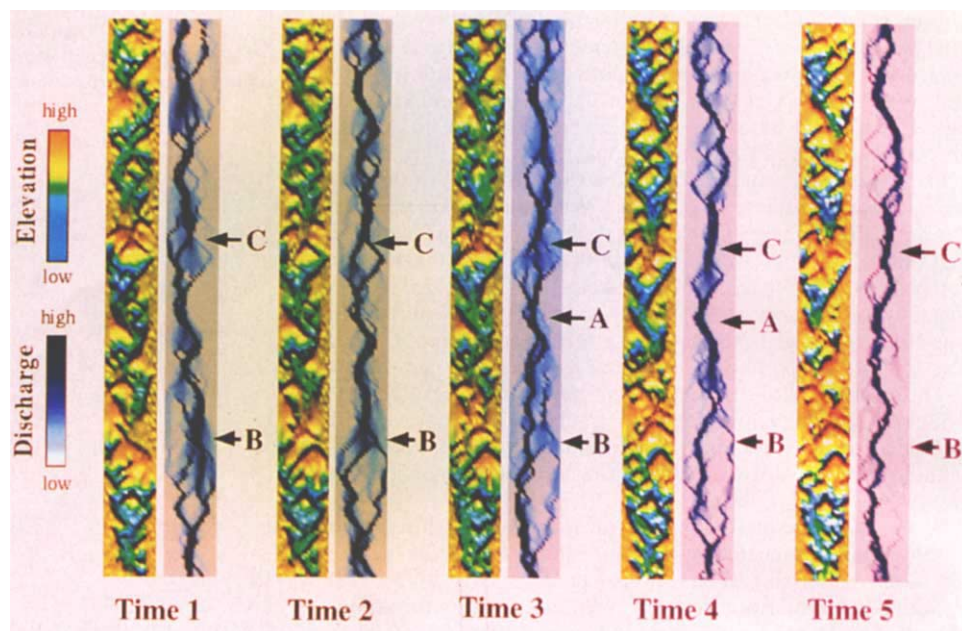
$$Q_{si} = K [Q_i S_i + \epsilon \sum Q_{i-1} S_{i-1}]^m \quad (4)$$

where Q_{si} is the amount of sediment transferred to downstream neighbour i from the cell in question, $m > 1$ (usually 2.5—see text), K is a constant adjusted so that the elevation changes by at most a few per cent of the mean elevation difference between rows in each time step, C is of the order of the average slope, $i - 1$ refers to contributions from upstream neighbours and $\epsilon \approx 0.3$. Black arrows represent lateral sediment transport, Q_{sij} (see text) given by: $Q_{sij} = F Q_{s0} S_{ij}$ where S_{ij} is the slope in lateral direction j , Q_{s0} is the total sediment transported out of the cell in question, and F is a constant adjusted so that Q_{sij} is of the order of a few per cent of Q_{s0} , consistent with observation^{15,16}.

embody this effect in a rule that transports small amounts of sediment down lateral slopes (see Box 1). This addition results in enhanced channel migration, but is not essential in producing a qualitatively realistic braided pattern. Moreover in the absence of the uphill-transport term, lateral transport always leads to damping of the initial perturbations.

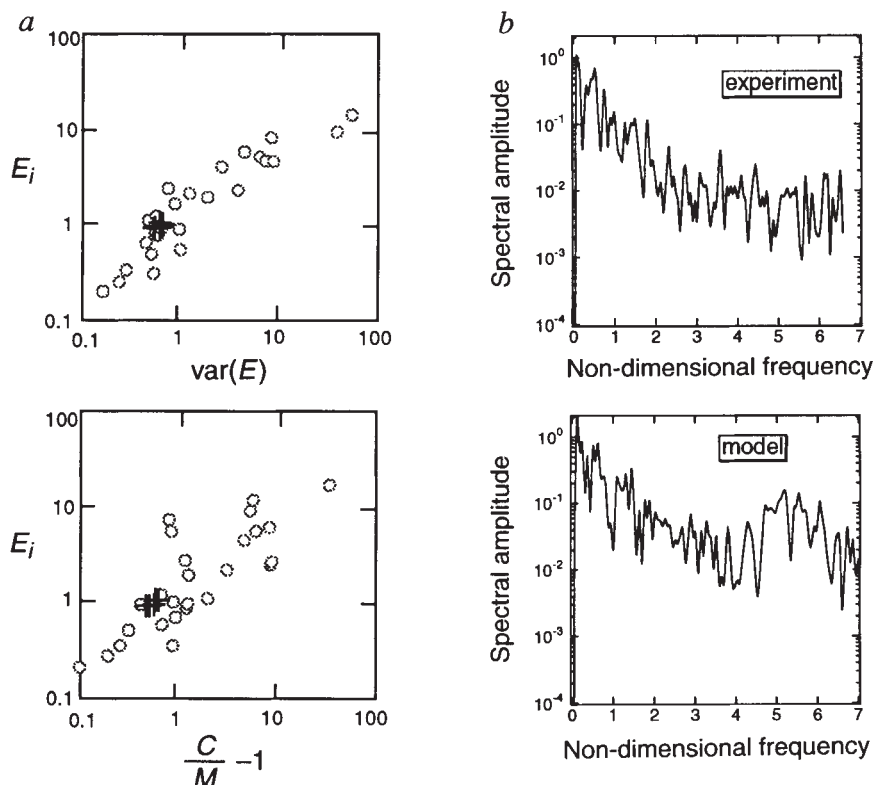
This model reproduces many of the phenomena observed in real braided streams, such as bar formation and reworking, channel splitting, channel switching and channel migration (Fig. 2). As in real rivers, the model pattern apparently remains statistically stable (Fig. 3a) although it never reaches a static steady state. It exhibits apparently unpredictable changes in configura-

FIG. 2 Images from successive times in one run of the model, showing 22×200 cells. Flow is from top to bottom. The left-hand image of each pair shows topography (minus the average slope), while the right-hand image shows discharge. The colour schemes for discharge and topography are shown at left. Topography is shown in shaded relief with lighting from the bottom (downstream). In this run we used: $n=1$ (see text and Box 1); sediment-transport rule 3 (see Box 1) with $m=2.5$ and $C=3$ times the average slope; and the lateral transport rule (see Box 1). The images are separated by 300 iterations. Channel migration occurs around location A, between times 3 and 4. A bar develops at location B by time 1, and the flow becomes shallow and develops multiple channels as it passes over the bar. By time 2, the bar has caused a split in the flow. By time 3, channel switching and new bar development just upstream of location B have cut off most of the water from the first bar. But by time 4 the new bar has caused another channel split, redirecting some flow back in the original direction. Then by time 5, the flow has reworked part of the new bar, cutting a channel across it at higher elevations than exist



in some areas around the bar, as we have observed in real braided streams in the laboratory. Bar development and reworking also occur at location C.

FIG. 3 Comparison of the cellular-model output with real braided rivers. *a*, Comparison of the averaged geometric statistics, after Howard *et al.*²⁹. E_i is the average number of channels per cross-section minus 1, C is the average stream width between the outermost banks, M is the average widest-channel width, and $\text{var}(E)$ is the variance of the number of channels. The plot shows how the overall widths of braided streams vary with the number of channels. Circles represent real rivers, and crosses show the model, measured at six different times during the run shown in Fig. 2. Such comparisons do not necessarily discriminate well between patterns generated by differing models²⁹; nonetheless, this comparison shows that the model remains statistically stable over time, and that its basic geometric statistics match those of real braided rivers. We are developing more sensitive tests of the model's realism. *b*, Power spectrum of a time series of the total sediment load of (top panel) a laboratory-model braided stream¹⁰, and (bottom panel) across a cross-section in our braided-stream (numerical) model. Both are broad-band 'brown' spectra with no obviously significant peaks. The frequency is nondimensionalized using a characteristic time $t_c = \bar{B}^2 \bar{\delta} / \bar{Q}_s$ where $\bar{B}$ is the average total channel width, $\bar{\delta}$ is the average standard deviation of the surface topography and $\bar{Q}_s$ is the average volumetric sediment flux. The spectral amplitude is normalized to integrate to unity. It has been suggested that the fluctuations develop because spreading of the discharge into many channels or unchannelled shallow areas decreases the overall sediment flux; migration of these zones causes the temporal



fluctuations^{21,30}. The sediment-load fluctuations in the model seem to arise in the same way.

tion indefinitely, even though no random influences are introduced after the run starts. A number of workers have reported self-organized ('autogenic') fluctuations or pulses in sediment flux in both laboratory and field braided streams^{18–22}. In our model, sediment load summed over a cross-section also fluctuates (Fig. 3b), and these fluctuations have broad-band spectra comparable to those of experimental bedload pulses (Fig. 3b).

We have done a series of runs in which we alter the width of the braid plain, keeping the average discharge per cell constant. As linear stability analyses^{23,24} predict, the number of channels initially formed increases linearly with the initial-flow width. Linear stability analyses address only this initial development. In our wide-channel runs, the flow eventually gathers into a smaller number of larger channels. If we start a run with the flow initially confined to a narrow trough with erodible banks, alternate bars and a meandering pattern develop first, with a wavelength of ~7–10 times the channel width, consistent with observation¹⁵. The meanders then erode the banks and grow in amplitude, until some channels cut across the point bars *en route* to a braided pattern, as has been observed in real braided streams²⁵.

The fundamental factors that distinguish braided from meandering rivers are still unclear. In our model, braiding develops when bedload transport leads to excess scour in flow convergences and deposition in divergences. This requires that the flow be sufficiently unconstrained laterally that it can change its width freely, and, if discharge or stream power is used to parametrize sediment flux, that the exponent in the sediment-flux law be >1 . The model results suggest that braiding is a simple, robust result of bedload transport under these conditions (although uphill and lateral sediment transport must also be included to produce realistic spatial patterns). It is difficult to compare this result directly with the various factors that have been proposed to discriminate between stream types in nature^{3,26}. It is, however, consistent with three known limiting cases; (1) networks formed by purely erosive processes are dendritic rather than braided²⁷, (2) braided rivers do not form in effectively cohesive materials where there is no bedload transport³, and (3) fully developed meandering cannot develop without some form of bank stabilization to constrain the flow laterally^{27,28}. This latter observation, together with the ubiquity of braiding in the model runs, suggests that braiding may be the fundamental instability of laterally unconstrained free-surface flow over cohesionless beds. We suggest that the meandering of rivers results from partial suppression of braiding instability by factors that either inhibit local redeposition, such as transport predominantly in suspension, or constrain the channel laterally, such as vegetation, cohesive sediment, or bedrock. □

Received 21 February; accepted 23 June 1994.

1. Ferguson, R.I. in *Braided Rivers* (eds Best, J. L. & Bristow, C. S.) 73–87 (Geological Soc. London, London, 1993).
2. Bristow, C. S. & Best, J. L. in *Braided Rivers* (eds Best, J. L. & Bristow, C. S.) 1–11 (Geological Soc. London, London, 1993).
3. Bridge, J. S. in *Braided Rivers* (eds Best, J. L. & Bristow, C. S.) 13–71 (Geological Soc. London, London, 1993).
4. Anderson, R. *Earth Sci. Rev.* **28**, 77–96 (1990).
5. Chase, C. G. *Geomorphology* **5**, 39–57 (1992).
6. Forrest, S. B. & Haff, P. K. *Science* **255**, 1240–1243 (1992).
7. Werner, B. T. & Hallet, B. *Nature* **361**, 142–145 (1993).
8. Tetzlaff, D. M. & Harbaugh, J. W. *Simulating Clastic Sedimentation* (Van Nostrand Reinhold, New York, 1989).
9. Werner, B. T. & Fink, T. M. *Science* **260**, 968–971 (1993).
10. Ashmore, P. E. *Process and Form in Gravel Braided Streams: Laboratory Modelling and Field Observations* (Univ. of Alberta, Edmonton, 1985).
11. Hong, L. B. & Davies, T. R. H. *Geol. Soc. Am. Bull. Part II* **90**, 1839–1859 (1979).
12. Ashmore, P. E. *Earth Surf. Processes & Landforms* **7**, 201–225 (1982).
13. Lysne, D. K. *J. hydr. Div. Am. Soc. Civil Engrs* **95**, 1835–1846 (1969).
14. Fredsoe, J. *J. fluid Mech.* **64**, 1–16 (1974).
15. Blondeaux, P. & Seminara, G. *J. fluid Mech.* **157**, 449–470 (1985).
16. Parker, G. *J. hydr. Engrng.* **110**, 197–199 (1984).
17. Parker, G. *J. fluid Mech.* **89**, 127–146 (1978).
18. Young, W. J. & Davies, T. R. H. *Earth Surf. Processes & Landforms* **16**, 499–511 (1991).
19. Hoey, T. *Prog. phys. Geogr.* **16**, 319–338 (1992).
20. Hoey, T. B. & Sutherland, A. J. *Earth Surf. Processes & Landforms* **16**, 447–462 (1991).
21. Ashmore, P. *Geogr. Annl.* **A73**, 37–52 (1991).
22. Goff, J. R. & Ashmore, P. *Earth Surf. Processes & Landforms* **19**, 195–212 (1994).

23. Parker, G. *J. fluid Mech.* **76**, 457–480 (1976).
24. Fredsoe, J. *J. fluid Mech.* **84**, 609–624 (1978).
25. Ashmore, P. E. *Can J. Earth Sci.* **28**, 326–341 (1991).
26. Robertson-Rintoul, M. S. E. & Richards, K. S. in *Braided Rivers* (eds Best, J. L. & Bristow, C. S.) 113–118 (Geological Soc. London, London, 1993).
27. Schumm, S. A., Mosley, M. P. & Weaver, W. E. *Experimental Fluvial Geomorphology* (Wiley, New York, 1987).
28. Schumm, S. A. & Khan, H. R. *Geol. Soc. Am. Bull.* **83**, 1755–1770 (1972).
29. Howard, A. D., Keetch, M. E. & Linwood, V. *Water Resour. Res.* **6**, 1647–1688 (1970).
30. Church, M. & Jones, D. in *Gravel-bed Rivers* (eds Hey, R. D., Bathurst, J. C. & Thorne, C. R.) 291–339 (J. Wiley, New York, 1982).

ACKNOWLEDGEMENTS. We thank P. Haff, G. Parker and R. Moeckel for helpful discussions, A. Howard and B. Werner for constructive comments on the manuscript, M. Person for use of computing facilities, and P. Jansen, D. Slawinski and P. Morin for technical help. This work was supported by the US National Science Foundation and the GEOFLUIDS programme.

Nd and Sr isotope evidence linking mid-ocean-ridge basalts and abyssal peridotites

Jonathan E. Snow, Stanley R. Hart & Henry J. B. Dick

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

PERIDOTITES found on the sea floor are widely believed to be residues left by mid-ocean-ridge basalt (MORB) melting. As such, their composition should provide insights into the nature of the sub-oceanic depleted mantle. But although these abyssal peridotites occur in mid-ocean ridge fracture zones^{1,2}, there is little other evidence in support of their genetic link with MORB, and doubts about it have been raised^{3–5}. Radiogenic isotopes should be able to provide a powerful test of the hypothesis, but previous studies^{3,8–10} on whole rocks have not provided unambiguous answers as they are generally altered by sea water. Here we present measurements of the Nd and Sr isotope compositions of a suite of abyssal peridotite clinopyroxenes which should have resisted alteration. Despite residual seawater contamination of Sr in the clinopyroxenes, the data demonstrate that the peridotites are from a depleted mantle source, identical to that of MORB. This provides a strong indication that abyssal peridotites are residues of MORB melting.

Most studies of abyssal peridotite either assume a relationship with MORB, and thus cannot be used to demonstrate its existence^{11,12}, or do not address this question at all^{13,14}. The best evidence so far for this link comes from worldwide covariations in basalt and abyssal peridotite major-element compositions^{15,16}.

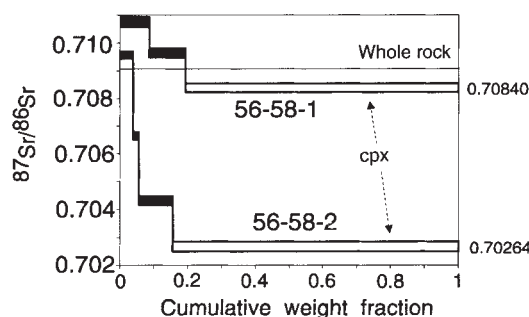


FIG. 1 Incremental leaching curve, sample 1011/76:56–58 (Islas Orcadas fracture zone). Dark boxes are leachate compositions. Open boxes labelled cpx are the residual clinopyroxenes after leaching, dissolved by standard means (see Table 1).

TABLE 1 Sample locations and isotope analyses

Sample locations						
Cruise	RC2709	RC2709	PS86	IO11/76	Vulc5	All-107-6†
Sample	6-2	25-139	6-37	56-58	41-29	40-35
Latitude	31° 54.9' S	32° 32.2' S	52° 21.0' S	54° 5.5' S	59° 5.2' S	54° 25.4' S
Longitude	57° 10.7' E	57° 3.8' E	13° 8.0' E	6° 17.1' E	16° 48.5' W	1° 34.2' E
Age (Myr)	0.6	14	1	9.7	2.3	17.0
Tectonic setting	Atlantis II fract. zone, N. RTI corner	All FZ, median ridge	12-15E zone: Dingaen FZ, NW Wall	Abyssal hills SE of Islas Orcadas FZ	59° S FZ, NW Wall	Bouvet FZ, NW Wall
Nd isotope ratios‡ (¹⁴³ Nd/ ¹⁴⁴ Nd) and concentrations§						
CPX	0.512994 (15)	0.513205 (14)	0.512960 (41)	0.513159 (33)	0.513073 (31)	0.513065 (62)
[Nd] ID	1.06	0.878	1.468	0.198	—	0.022
[Nd] SIMS	1.8	—	2.37	—	0.83	0.06
Sr isotope ratios (⁸⁷ Sr/ ⁸⁶ Sr) and concentrations						
CPX	0.704870 (13)	0.703173 (23)	0.703350 (23)	0.702644 (25)	0.702902 (22)	0.702869 (75)
Weight (mg)	90.71	65.00	81.12	67.72	55.45	73.82
[Sr] ID	24.1	0.64	6.42	0.15	1.34	0.05
[Sr] SIMS	30	—	5.40	—	1.29	0.87
Whole rock	0.70622 (23)	—	0.709141 (27)	0.709100 (21)	0.709448 (70)	0.709041 (489)
[Sr] ID	4.13	—	311.5	6.85	20.7	3.52
Leach 1	0.706929 (34)	0.709003 (23)	0.708800 (85)	0.709579 (28)	0.709116 (169)	0.709039 (24)
Weight (mg)	4.63	7.73	9.81	3.21	3.94	4.77
[Sr] ID	17.9	32.48	1105	6.72	7.74	9.92
Leach 2	0.705283 (574)	0.707700 (21)	0.709094 (26)	0.706716 (123)	—	0.706703 (45)
Weight (mg)	1.87	7.28	6.84	1.25	1.48	2.46
[Sr] ID	37.1	4.60	168.6	4.42	0.62	2.38
Leach 3	0.705118 (123)	0.705657 (24)	0.707994 (24)	0.704374 (38)	0.702776 (232)	0.704421 (30)
Weight (mg)	10.17	6.22	13.71	8.40	6.98	12.31
[Sr] ID	35.7	2.53	29.63	0.59	0.84	0.48

Abbreviations used: FZ, fracture zone; All, Atlantis II; PS, Polar Stern; VULC, Vulcan; RC, Robert Conrad; IO, Islas Orcadas; RTI, ridge transform intersection. ID, isotope dilution; SIMS, secondary ionization mass spectrometry. 2–10 kg of bulk peridotite was crushed to ~400–800 µm and washed in deionized water and ethanol. Repeated magnetic separation resulted in a few hundred milligrams of clinopyroxene-rich (~40%) material. Clinopyroxene fragments were hand-picked to yield 100–150 mg of nearly pure (~95%) clinopyroxene. This separate was then crushed again, sieved to 200–400 µm size, washed again and hand-picked again. Only clear clinopyroxene grains were selected, free of inclusions of spinel, serpentine and clay, and showing good cleavage. Three leaches were made on each sample, using a leaching solution of 6.2 N HCl, 5% HF. The first leach was cold for 5 min, the second in an ultrasonic bath for 10 min and the third for 10 min in the ultrasonic bath followed by 10 min at 125 °C. The samples were dried and weighed between leaches. The weights dissolved are given in mg. The final residue was dissolved in 4:1 HF:HClO₃ and analysed by standard methods.

* Crustal ages from location, plate geometry and spreading velocities.

† Nd and Sr were measured on different leached separates of this sample. Nd analysis was carried out at CRPG, Nancy, France, normalized to standard compositions stated below.

‡ ¹⁴³Nd/¹⁴⁴Nd collected in multi-dynamic mode, normalized to ¹⁴⁶Nd/¹⁴⁴Nd of 0.7219. The average value of LaJolla Nd isotope composition standard measured during the course of this work is 0.511833 ± 19 (1σ, N = 19).

§ All concentrations are determined by isotope dilution (except for SIMS results as noted), given in p.p.m., and bear a nominal analytical uncertainty of 1%.

|| ⁸⁷Sr/⁸⁶Sr collected in multi-dynamic mode, normalized to ⁸⁶Sr/⁸⁸Sr = 0.1194 and corrected to NBS SRM 987 Sr isotope composition standard value of 0.710240. The uncorrected average NBS 987 value is 0.710323 ± 20 (1σ, N = 16).

Radiogenic isotope systems (Sr, Nd, Pb and Os) provide a powerful test of a genetic link between rock series. They are not affected by melting and crystallization (as are elemental concentrations and ratios), nor do they rely on a particular genetic model for validity (for example, near-fractional versus batch melting). The ratio of radiogenic to non-radiogenic isotopes (for example, ¹⁴³Nd/¹⁴⁴Nd) provides a 'genetic fingerprint' for different source regions, because melt and residue equilibrate during melting and thus have identical isotopic compositions. Attempts to evaluate the relationship between MORB and abyssal peridotite using isotope systematics of MORB and peridotite whole rocks have produced either ambiguous results^{6,7} or ones that seemed to contradict it^{3,8–10}.

Currently, the only reliable isotopic data from near-mid-ocean-ridge mantle rocks are those from the St. Peter-Paul islets^{6,7}. These small outcrops of peridotite are situated in the Atlantic Ocean, and occupy a tectonic setting similar to those of abyssal peridotites, but are compositionally dissimilar. They may represent either ocean island basalt (OIB)⁶, or continental⁷ mantle and differ from abyssal peridotites in many ways. Isotopically they resemble OIB and thus greatly confuse the issue surrounding the MORB–abyssal peridotite genetic link.

Further, isotopic studies in peridotites are complicated by the difficulty of the analysis. Abyssal peridotites are rare compared to MORB, and those containing fresh clinopyroxene rarer still. Almost all abyssal peridotites have undergone extensive serpentinization and weathering, resulting in profound changes in elemental abundances and isotopic ratios^{17–19}. Also, most radiogenic elements are found in very low concentrations in abyssal peridotite clinopyroxenes, even compared to other classes of peridotite¹¹. Lastly, clinopyroxene has a low abundance in abyssal peridotites, rarely making up >10% of the rock^{15,16}.

To test the MORB–abyssal peridotite relationship, we analysed a suite of 6 separated and leached abyssal peridotite clinopyroxenes from the southwest Indian and American–Antarctic ridges (Table 1). The rock samples selected are all spinel harzburgites and lherzolites, and have 3–12% model clinopyroxene¹⁸. All are serpentinized (60–90%), but do not contain high-temperature alteration phases or plagioclase. Sr and Nd isotope compositions and concentrations in clinopyroxenes and leachates were measured by thermal ionization mass spectrometry.

Figure 1 shows leaching profiles for sample IO11/76:56–58. Two attempts at leaching this sample resulted in greatly different

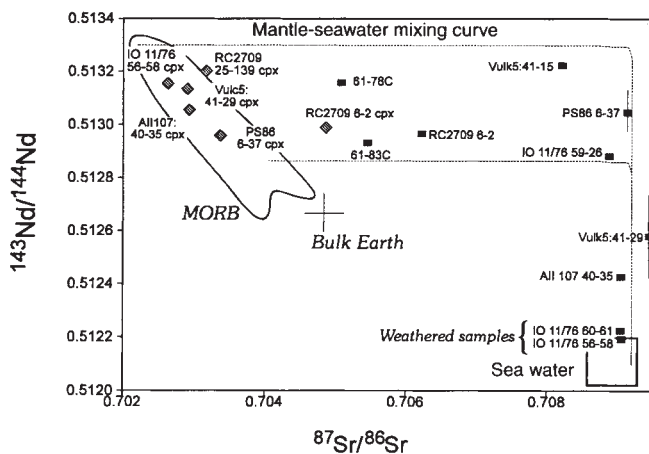


FIG. 2 $^{143}\text{Nd}/^{144}\text{Nd}$ vs. $^{87}\text{Sr}/^{86}\text{Sr}$ in abyssal peridotite whole rocks and clinopyroxenes. Broken lines enclosing the data points are calculated peridotite-seawater mixing curves. Diamonds denote clinopyroxenes, while rectangles denote whole rocks¹⁹. Sample labels refer to analyses in Table 1.

Sr isotope compositions. In general, replicate leaching experiments rarely produced agreement to better than several times the analytical uncertainty, so leaching alone cannot remove all contaminant Sr. All likely contaminants have a higher $^{87}\text{Sr}/^{86}\text{Sr}$ (sea water is 0.7092, and some orphan isotopic compositions reach 0.726 (ref. 19)). The measured Sr isotope ratios thus appear to be mixtures of mantle and seawater values. They provide quantitative information about the mantle in that they provide an upper limit on the primary Sr isotopic composition of the clinopyroxene. The lowest Sr isotope compositions show that they were derived from the depleted mantle, even though the exact composition cannot be determined. We cannot rule out the possibility that the higher Sr isotope compositions (for example, RC2709:6-2) could reflect mantle values. These isotopic compositions would then imply a high $^{143}\text{Nd}/^{144}\text{Nd}$, high $^{87}\text{Sr}/^{86}\text{Sr}$ component, as yet unknown. Also, the whole rock data define an Nd-Sr array that follows a mantle-seawater mixing curve (Fig. 2). Seawater contamination thus seems more likely.

The Sr isotope analyses can be used to constrain the quality of the Nd isotope analyses. Figure 2 shows Nd and Sr isotope compositions of the clinopyroxenes and whole rocks, along with seawater mixing curves reflecting different possible mantle compositions. For samples where the residual contamination is small enough for the Sr isotope composition to be close to a mantle value, then the effect on the Nd isotope composition is substantially smaller than the analytical precision. For instance, assuming that the true Sr isotope composition of sample RC2709:6-2 lies in the mantle field, the change in $^{143}\text{Nd}/^{144}\text{Nd}$ is $<2 \times 10^{-8}$, and is substantially less in the other samples. This reasoning does not apply to All107:40-35, where Sr and Nd isotopes were measured on different separates.

The Nd (and to some extent Sr) isotope ratios in clinopyroxenes agree generally with their values in MORB (Figs 2 and 3). Both are derived from mantle that has been depleted over geological time in large-ion lithophile elements. These observations provide the best evidence to date showing that abyssal peridotites are the residues of MORB melting, consistent with their tectonic setting and major element covariation^{15,16}.

The peridotite clinopyroxene isotopic compositions as a group do not define a highly depleted mantle endmember composition, as has been postulated to exist beneath mid-ocean ridges^{1,20}. Rather than residing at the depleted end of the MORB field, they cover almost the entire range of MORB isotopic compositions (Figs 2 and 3). This might suggest that mantle sources are frequently intermediate in composition to the mantle

endmember extremes^{6,21}. The data overlap somewhat with the field for OIB, but a significant fraction of them lies outside the range of OIB compositions, particularly Indian Ocean OIBs (see Fig. 3). An OIB source thus cannot account for the isotopic compositions of abyssal peridotites. Arguably, there might be an OIB component to some of these compositions, but this case is indistinguishable from the one above^{6,21}.

By comparing the abyssal peridotite data with those from basalts from the same region, we can draw some tentative conclusions about melt generation beneath mid-ocean ridges. Two samples came from the northern rift of the Atlantis II fracture zone (Fig. 3) and bracket the $^{143}\text{Nd}/^{144}\text{Nd}$ of nearby basalts. They thus exhibit a greater range of variability than over 80 of the basalt and gabbro analyses from this relatively homogeneous ridge segment^{18,22,23}.

The clinopyroxene analysed from both the 12–15° E disturbed region and the 59° S fracture zone agree with basalts from the same regions²⁴ whereas those from the Islas Orcadas fracture zone and the nearby Bouvet fracture zone are more depleted than the nearby basalts (see Fig. 3).

The existence of regions in which the peridotites appear more isotopically extreme than the basalts (Atlantic II North and Bouvet) suggests that MORB reflects small-volume near-fractional melts^{11,25}. A large-volume batch process would probably homogenize such isotopic heterogeneity²⁶. This isotopic diversity echoes that seen at Ronda, although not nearly to the same degree²⁷. It is however difficult to compare the results from Ronda to those from the mid-ocean-ridge mantle due to the great differences in their tectonic setting, age and composition.

Vein melting is a means of explaining an apparent link between isotopic composition and degree of partial melting in

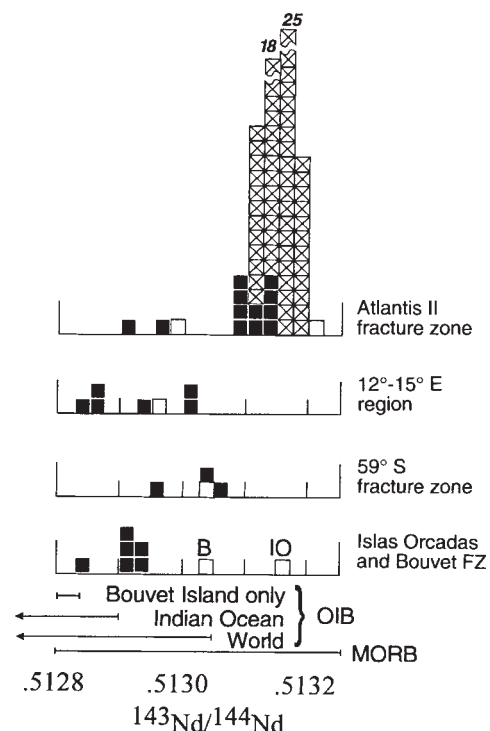


FIG. 3 Peridotite, basalt and gabbro isotopic compositions. Localities are listed at the right. Peridotites are shown as open blocks, gabbros as crossed blocks and basalts (MORB) as black blocks. Basalt data from refs 24, 29, 30. Ranges of data for worldwide MORB, worldwide OIB and Indian Ocean OIB are also shown. The width of each bin is approximately the same as the 2σ uncertainty of the peridotite analyses.

some basalts²⁸. Enriched veins in the mantle source have a lower melting point, and are sampled by the first fractions of melting. This results in (relatively) enriched melts in the early fractions, followed by more depleted melts. Residues of melting, having lost their vein components, would then have more depleted isotopic compositions than basalts derived from both early and late melts. This may explain the depleted Nd ratios in the Atlantis II and Bouvet-Islands Orcadas regions.

The relatively enriched sample from the Atlantis II fracture zone (RC2709:6-2) comes from a sample containing a clinopy-

roxene vein and textural evidence for reaction of the sample with an infiltrating liquid (probably a silicate melt) within the mantle¹⁸. Lower ¹⁴³Nd/¹⁴⁴Nd in this sample than in basalts from the region seems to support vein melting by showing that enriched veins exist in mid-ocean-ridge mantle. However, peridotites that had risen through a melting region should be devoid of such veins. Thus a late-stage melt infiltration event high in the lithosphere is a more likely interpretation of this sample, although the origin of this infiltrating fluid is unknown. □

Received 7 April; accepted 1 August 1994.

- Miyashiro, A., Shido, F. & Ewing, M. *Contr. Miner. Petrol.* **23**, 117-127 (1969).
- Aumento, F. & Loubat, H. *Can. J. Earth Sci.* **8**, 631-663 (1970).
- Bonatti, E. *J. geophys. Res.* **76**, 3825-3831 (1971).
- Coleman, R. G. *Ophiolites* (Springer, 1977).
- Salters, V. & Shimizu, N. *Geochim. cosmochim. Acta* **52**, 2177-2182 (1990).
- Roden, M., Hart, S., Frey, F. & Melson, W. *Contr. Miner. Petrol.* **85**, 376-398 (1984).
- Bonatti, E. *Nature* **345**, 800-802 (1990).
- Bonatti, E., Honnorez, J. & Ferrara, G. *Earth. planet. Sci. Lett.* **9**, 247-256 (1970).
- Hart, S. R. *Carnegie Inst. Wash. Yk* **71**, 290-292 (1972).
- Roy-Barman, M. thesis, Univ. Paris 7 (1993).
- Johnson, K. T. M., Dick, H. J. B. & Shimizu, N. *J. geophys. Res.* **95**, 2661-2678 (1990).
- Elthon, D. *J. geophys. Res.* **97**, 9015-9025 (1992).
- Shibata, T. & Thompson, G. *Contr. Miner. Petrol.* **93**, 144-159 (1986).
- Cannat, M., Bideau, D. & Bougault, H. *Earth planet. Sci. Lett.* **109**, 87-106 (1992).
- Dick, H. J. B., Fisher, R. & Bryan, W. B. *Earth planet. Sci. Lett.* **69**, 88-106 (1984).
- Michael, P. & Bonatti, E. *Earth planet. Sci. Lett.* **73**, 91-104 (1985).

- Kimball, K. L., Spear, F. S. & Dick, H. J. B. *Contr. Miner. Petrol.* **91**, 307-320 (1985).
- Snow, J. E. thesis, MIT Woods Hole Oceanogr. Inst. (1993).
- Snow, J. E., Hart, S. R. & Dick, H. J. B. *Science* **262**, 1861-1863 (1993).
- Zindler, A. & Hart, S. R. *A. Rev. Earth Planet. Sci.* **14**, 493-571 (1986).
- Barling, J. & Goldstein, S. L. *Nature* **348**, 59-62 (1990).
- Snow, J. E., Dick, H. J. B. & Hart, S. R. *EOS* **68**, 408 (1987).
- Kempton, P. D., Hawkesworth, C. J. & Fowler, M. *Proc. ODP Sci. Res.* **118**, 127-142 (1991).
- le Roex, A., Dick, H. J. B. & Watkins, R. T. *Contr. Miner. Petrol.* **110**, 253-268 (1992).
- Sobolev, A. & Shimizu, N. *Nature* **363**, 151-154 (1993).
- Hofmann, A. W. & Hart, S. R. *Earth planet. Sci. Lett.* **38**, 44-62 (1978).
- Reisberg, L. & Zindler, A. *Earth planet. Sci. Lett.* **81**, 29-45 (1986).
- Zindler, A., Staudigel, H. & Batiza, R. *Earth planet. Sci. Lett.* **70**, 175-195 (1984).
- le Roex, A., Dick, H. J. B., Erlank, A., Reid, A. & Frey, F. *Contr. Miner. Petrol.* **90**, 367-380 (1985).
- le Roux, A. et al. *J. Petrol.* **24**, 267-318 (1983).

ACKNOWLEDGEMENTS. We thank M. Kurz for discussions and L. Reisberg, P. Michael and J. A. Snow for reviews. This study was supported by an NSF graduate fellowship and a WHOI Ocean Ventures Fund grant to J.E.S., and by US N.S.F. grants to S.R.H. and H.J.B.D.

Coupling of tree transpiration to atmospheric turbulence

D. Y. Hollinger^{*†}, F. M. Kelliher^{*}, E.-D. Schulze[‡] & B. M. M. Köstner[‡]

^{*} Manaaki Whenua-Landcare Research, Box 31-011, Christchurch, New Zealand

[‡] Lehrstuhl Pflanzenökologie der Universität Bayreuth, Box 10 12 51, 8580 Bayreuth, Germany

UNDERSTANDING mass and energy exchange between vegetation and the atmosphere is essential for determining the future state of the climate system^{1,2} and responses of plant communities. Plant water use is at present described by steady-state transport models^{3,4}, even though transport in the boundary layer is turbulent^{5,6}. This is especially true for forests, where the canopy air space is a chaotic environment where large turbulent events alternate with smaller-scale mixing^{3,4,7}. Here we demonstrate that the turbulent nature of the atmosphere affects plant processes. We propose that the dynamic nature of plant-atmosphere coupling represents a previously unrecognized feedback that influences plant water use and transport.

Plant water use (transpiration, E) is regulated by the available energy (R_n) and air saturation deficit (D) above the canopy (Fig. 1a). The relative importance of these two factors in regulating plant or ecosystem water use is theoretically summarized in a decoupling coefficient, Ω , ($0 \leq \Omega \leq 1$) derived from the Penman-Monteith equation^{3,4,8}. When $\Omega = 0$, transpiring plant leaves and the air in the canopy are completely coupled to the atmosphere overhead, and E is completely coupled to D . When $\Omega = 1$, leaves are decoupled or aerodynamically isolated from the atmosphere and E is determined by R_n . Because tree canopies are tall, they generate more turbulence than shorter vegetation, increasing the coupling of E to D . For example, half-hourly Ω values in our 30-m-tall beech forest were generally <0.25 (ref. 9).

We find that high frequency fluctuations are superimposed upon the daily pattern of sap flux or transpiration (Fig. 1a). A comparison of tree sap flux and overhead atmospheric condi-

tions suggests that these fluctuations in transpiration are related to the intermittent nature of atmospheric transport (Fig. 1b). Two displacement events marked by the variation in horizontal windspeed (u), are seen in this 6-min record. At the beginning of each event, u increases rapidly, and cooler, drier air moves down into the canopy. In the quiescent period between events, air temperature (T) and specific humidity (q) gradually increase. The flux of water up a tree stem increases at the start of displacement events and moves inversely to variations in atmospheric specific humidity.

Spectral analysis shows that the relative magnitude of fluctuations of sap flow in the frequency domain are similar to those of water vapour flux in the atmosphere above the tree (Fig. 2a). These power spectra reach a maximum at ~ 0.003 - 0.004 Hz. According to Taylor's frozen field hypothesis, this corresponds to a length scale of the dominant turbulent eddies of about 200-300 m at the measurement windspeed, in agreement with theoretical expectations of ~ 8 times the vegetation height¹⁰. We also examined the coherence (correlation in the frequency domain) between variations in tree sap flux and atmospheric conditions. Because changes in sap flow may lag atmospheric changes, we calculated relative coherences (summed over the frequency range of 10^{-3} to 1 Hz) as a function of the time lag (Fig. 2b). We examined separately the specific humidity and air temperature components of D . This analysis suggests that fluctuations in sap flow at 15 m above the ground (F_{top}) lagged variations in q by ~ 1 s and changes in T by ~ 12 s. Sap flux measured at the bottom of a tree (F_{base}) lagged variations in q and T by ~ 13 and ~ 22 s, respectively, and lagged fluctuations in F_{top} of the same tree by ~ 6 s. Variations in T tended to lag fluctuations in q by several seconds. The lag between F_{top} and F_{base} suggests that water is transiently drawn from a supply within the stem. However, the dynamic stem capacitance represented by 6 s of flow is minimal (<0.1 dm³ in a 0.8-m-diameter tree).

The difference between the lags of sap flux and variations in T and q is intriguing, and may pertain to how these factors influence transpiration. For a leaf, transpiration is the product of the stomatal conductance (g) and the difference in specific humidity between the leaf (q_l^*) and the air adjacent to the leaf surface (q_a): $E = g(q_l^* - q_a)$.

When drier air moves into a canopy, E immediately increases because q_a has decreased. Stomatal response may take several

[†] Present address: USDA Forest Service, NE Exp. Station, PO Box 640, Durham, NH 03824, USA.

TABLE 1 Correlation of sap flux variations between trees of the same or different strata in an undisturbed forest

	<i>n</i> *	Mean†	δ †	% Sig.‡	% Neg.§
Among emergents	3	0.281	0.104	100	0
Between emergents and canopy	15	0.158	0.131	66.7	0
Between emergents and subcanopy	18	0.042	0.149	33.3	5.6
Among canopy	10	0.063	0.102	30.0	10
Between canopy and subcanopy	30	0.053	0.118	23.3	6.7
Among subcanopy	15	-0.038	0.209	26.7	26.7

* Number of possible correlations between or among three strata of trees in a forest patch consisting of 3 emergent, 5 canopy and 6 subcanopy trees. Correlations were calculated from 240 1-min residual clear-day sap flux values centred on solar noon. The general trend of sap flux was removed by subtracting a best-fit low-order polynomial from the raw data record of each tree.

† Mean and standard deviation of the *n* correlation coefficients possible among or between tree strata.

‡ Per cent of the *n* correlations that are positive and significant ($P < 0.05$).

§ Per cent of the *n* correlations that are negative and significant ($P < 0.05$).

minutes¹¹. Fluctuations in *T* can affect *E* by changing leaf temperature and hence q_i^* , as well as q_a . However, because of the foliage heat capacity, some seconds are required for a change in leaf temperature to occur. A dynamic energy balance calculation¹², suggests that the $1 - e^{-1}$ time constant for a temperature change in beech leaves ranges between about 5 and 15 seconds, depending on the windspeed. Our difference between $F_{top}:q$ and $F_{top}:T$ lags of ~ 10 s is compatible with these estimates.

Atmospheric specific humidity and temperature fluctuations and lagged tree transpiration are coherent over a range of fre-

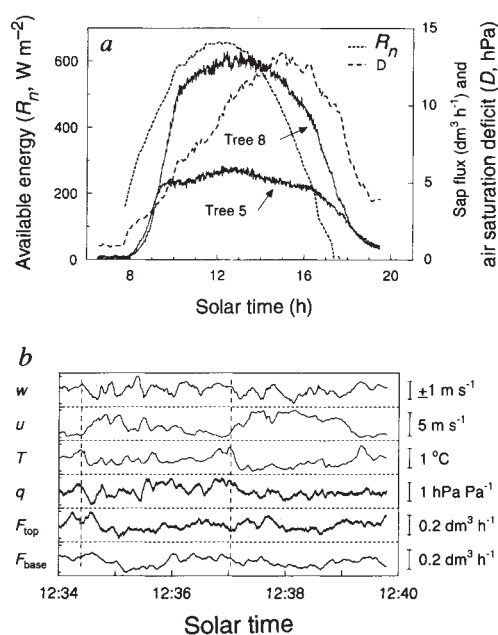


FIG. 1 a, Typical clear-day pattern of available energy (R_n), atmospheric saturation deficit (D), and the flux of water up two adjacent tree stems in a 30-m-tall *Nothofagus* forest in New Zealand (12 April 1991). Sap flux was measured in 14 trees in a patch of forest with a heated-stem, mass flow method⁹ at a height of 1.5 m. Available energy and air saturation deficit were measured from a tower that extended 6 m above the forest. b, High-frequency (10 Hz) recording of vertical wind speed (w), horizontal wind speed (u), air temperature (T) and specific humidity (q) measured at 36 m. The start of two displacement events is marked by the dashed vertical lines. Also shown is the sap flux at the top of the stem at 15 m (F_{top}) and 1.5 m (F_{base}) for a 32-m-tall emergent tree adjacent to the measurement tower and parallel with the wind. The micrometeorological instrumentation and site is described in refs 9, 21 and 22.

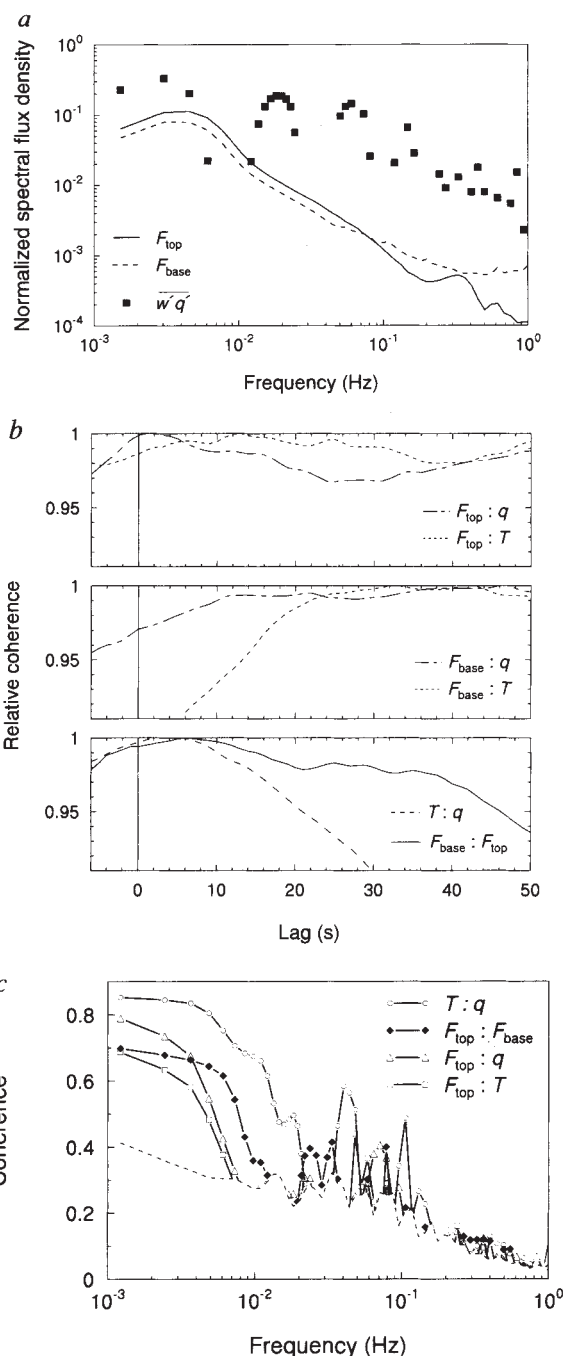


FIG. 2 a, Cospectra of w/q' (water vapour flux at 36 m) and power spectra of stem sap flux at 15 m and 1.5 m. Power and coherence spectra were computed by fast Fourier transform²³ from series of 8,192 data points that had been recorded at 10 Hz. Coordinates were rotated in two-dimensions to force the mean vertical (w) and lateral (v) wind velocities to zero. Each time series was tapered with a cosine window and linear trends were removed. The raw estimates were logarithmically blocked to produce individual spectra that were normalized by their variance. Four spectra were averaged by frequency to produce the final spectra. b, Relative coherences (summed over the interval of 0.001 and 1 Hz) between q , T , F_{top} and F_{base} as a function of the time lag between the data arrays. The lag values stated in the text were determined by linear extrapolation of the relative coherence-lag relationship up to a relative coherence value of 1 beginning at a lag of -6 s. Data collected from $\sim 13:00$ – $15:00$ on a sunny afternoon, 11 April 1991. c, Coherence between F_{top} and F_{base} , q and T , after accounting for the appropriate lag (compare b). Also shown is the coherence between T and q . To estimate the significance of our coherence values, we calculated the coherence between F_{top} and a number of series constructed of white noise. Ninety-five per cent of the F_{top} -noise coherence values lie below the dashed line. Coherence values below that line are not shown.

quencies (Fig. 2c). Fluctuations in sap flow at the top and base of a tree are coherent for frequencies less than 0.01 Hz, but at higher frequencies this coherence declines significantly because the tree stem acts as a low-pass filter. A consequence of the coherence between tree transpiration and large-scale turbulent fluctuations of q and T , is that as eddies sweep across a forest, sap flux varies in a correlated manner among neighbouring trees (Table 1), creating an invisible analogue to fields of grain rippling in the wind. The correlation coefficients decrease, however, between and among trees lower in the canopy, suggesting that most turbulent gusts do not penetrate through the tree canopy¹³. This demonstrates one way that forest structure affects Ω and evaporation.

We have demonstrated that the impact of atmospheric turbulence extends across the leaf-air boundary layer to affect the mass transport of water in trees. This finding contrasts with diffusive or active transport models of water flow in plants¹⁴. There are several important implications of a chaotic component to water transport in the soil-plant-atmosphere continuum. Because stomata respond to changes in transpiration^{15,16} and this nonlinear response is both asymmetrical (closing is more rapid than opening) and overlaps the frequency of displacement events¹¹, the intensity and frequency of atmospheric turbulence may provide a previously unrecognized source of feedback that affects stomata and plant functioning. High-frequency variation in transpiration flux will lead to a similar variation in xylem pressure potential that may potentially influence protein synthesis in the roots or at other sites remote from the foliage. Hydraulic pressure fluctuations have been linked with plant electrical action potentials and these action potentials with changes in

protein production¹⁷⁻¹⁹. Plants respond asymmetrically and non-linearly to high-frequency variations in light²⁰. We regard the turbulent nature of atmospheric transport to be as important for plant processes as the response to fluctuating light. □

Received 10 May 1993; revised 16 May 1994; accepted 18 July 1994.

- Shukla, J., Nobre, C. & Sellers, P. *Science* **247**, 1322-1325 (1990).
- Bonan, G. B., Pollard, D. & Thompson, S. L. *Nature* **359**, 716-718 (1992).
- Monteith, J. L. *Soc. exp. Biol.* **19**, 205-234 (1965).
- McNaughton, K. G. & Jarvis, P. G. in *Water Deficits and Plant Growth* Vol. VII (ed. Kozlowski, T. T.) 1-43 (Academic, New York, 1983).
- Gao, W., Shaw, R. H. & Paw, U. K. T. *Boundary-Layer Meteorol.* **47**, 349-377 (1989).
- Bergström, H. & Höglström, U. *Boundary-Layer Meteorol.* **49**, 231-263 (1989).
- Finnigan, J. J. & Raupach, M. R. in *Stomatal Function* (eds Zeiger, E., Farquhar, G. D. & Cowan, I. R.) 385-430 (Stanford University Press, Stanford, 1987).
- Martin, P. *Agric. For. Met.* **49**, 45-53 (1989).
- Köstner, B. M. M. *et al. Oecologia* **91**, 350-359 (1992).
- Raupach, M. R., Finnigan, J. J. & Brunet, Y. *Fourth Australasian Conference on Heat and Mass Transfer* 75-90 (Christchurch, New Zealand, 1989).
- Grant, D. A. & Zeiger, E. *Plant Physiol.* **81**, 865-868 (1986).
- Gates, D. M. *Biophysical Ecology* (Springer, New York, 1980).
- Denmead, O. T. & Bradley, E. F. in *The Forest-Atmosphere Interaction* (eds Hutchinson, B. A. & Hicks, B. B.) 421-442 (Reichel, Dordrecht, Holland, 1985).
- Zimmermann, U. *et al. Phil. Trans. R. Soc. Lond. B* **341**, 19-31 (1993).
- Nonami, H., Schulze, E.-D. & Ziegler, H. *Planta* **183**, 57-64 (1990).
- Mott, K. A. & Parkhurst, D. F. *Plant Cell Environ.* **14**, 509-515 (1991).
- Malone, M. & Stankovic, B. *Plant Cell Environ.* **14**, 431-436 (1991).
- Malone, M. *Planta* **187**, 505-510 (1992).
- Wildon, D. C. *et al. Nature* **360**, 62-65 (1992).
- Pearcy, R. W. & Pflüsch, W. A. in *Ecophysiology of Photosynthesis* (eds Schulze, E.-D. & Caldwell, M. M.) 343-359 (Springer, Berlin, 1994).
- Kelliher, F. M. *et al. Agric. and For. Met.* **62**, 53-73 (1992).
- Hollinger, D. Y. *et al. Ecology* **75**, 134-150 (1994).
- Carter, G. C. & Ferrie, J. F. in *Programs for Digital Signal Processing* (eds Digital Processing Committee) 2.3-1 to 2.3-18 (IEEE Press, New York, 1979).

ACKNOWLEDGEMENTS. This work was supported by funding from the New Zealand Foundation for Research, Science and Technology, New Zealand Department of Conservation, Bundesministerium für Ernährung, Landwirtschaft und Forsten, and the Bundesminister für Forschung und Technologie. We thank J. Byers, J. Hunt, T. McSeveny, R. Meserth and P. Weir for technical assistance, and J. Orwin and D. Whitehead for their comments on the manuscript.

***Sordes pilosus* and the nature of the pterosaur flight apparatus**

David M. Unwin* & Natasha N. Bakhurina††

* Department of Geology, University of Bristol, Queen's Road, Bristol BS8 1RJ, UK

† Palaeontological Institute, Russian Academy of Sciences, Profsoyuznaya 123, Moscow 117647, Russia

It is now generally accepted that pterosaurs, Mesozoic reptiles, were true fliers, but the nature of their flight apparatus is still much disputed. Evidence has been presented in favour of bird-like reconstructions with narrow, stiff wings free of the legs¹⁻⁶ and bat-like reconstructions with extensive wings incorporating both fore and hind limbs⁷⁻¹⁰, but the Solnhofen Limestone pterosaurs, upon which these models are based, are not sufficiently well preserved to resolve these conflicting interpretations. Here we present a new model, founded on *Sordes pilosus* from the Jurassic of middle Asia (ref. 11, and N.N.B. and D.M.U., manuscript submitted), in which exceptionally well preserved wing membranes show that the hind limbs of pterosaurs were intimately involved in the flight apparatus; connected externally to the main wing membrane and internally by a uropatagium, controlled by the fifth toe. *Sordes* also reveals that, uniquely among flying vertebrates, pterosaurs had a structurally non-homogenous flight surface with a stiffened outer half and a softer, more extensible inner region.

The Upper Jurassic Karabastau formation (Oxfordian-Kimmeridgian) of Karatau in Kazakhstan¹², a sequence containing very fine-grained lake deposits, has yielded many exception-

ally well preserved fossils¹³, including insects, fish, turtles, a salamander, lizard, crocodile and two pterosaurs: *Batrachognathus volans*, a small insectivore^{14,15}, and *Sordes pilosus*, a long-tailed piscivore with a wingspan of 0.65 m (ref. 11, and N.N.B. and D.M.U., manuscript submitted). *Sordes* is represented by seven specimens, varying in completeness, but often naturally articulated, and exhibiting little evidence of post-mortem disturbance. The wing membranes, preserved as high-fidelity impressions picked out by dark, mineralized soft tissues (Fig. 1a) are complete, their edges easily traceable (Fig. 1b), and, in some cases, very fine internal structures such as the individual strands of wing fibres, each less than 10 µm in diameter, are clearly visible (Fig. 2a, b). Wing preservation is substantially better than in Solnhofen pterosaurs, but three-dimensional structures, as reported in a pterosaur wing from the Cretaceous of Brazil¹⁶ have not been found.

As in other pterosaurs, there is a small, triangular propatagium located in the angle of the humerus and forearm, and manipulated by a wrist bone, the pteroid (Fig. 3). The main wing membrane (cheiropatagium) is attached to the rear edge of the forelimb, to the body wall from the shoulder to the hip and to the anterior margin of the hind limb as far as the ankle (Fig. 1a, c). Attachment of the cheiropatagium to the leg has also been reported in *Pterodactylus*⁹, but the exact extent, whether to the knee or to the ankle, is disputed. Some^{4,6} restrict it to the thigh, but others^{7,9} argue for a more extensive attachment, an idea supported by two examples of *Pterodactylus*^{17,18}, in which impressions of the cheiropatagium are preserved adjacent to the lower leg. Elsewhere, wing impressions are rather ambiguous, or not preserved at all, but morphometric studies¹⁹ have revealed a strong correlation between wing and leg length in pterosaurs. A similar correlation occurs in bats where the fore and hind limbs are connected by the flight apparatus, but not in birds in which these structures are physically and functionally independent. In the absence of evidence to the contrary we conclude that attachment of the cheiropatagium to the hind limb was probably universal in pterosaurs.

† Present address: Department of Geology, University of Bristol, Queen's Road, Bristol BS8 1RJ, UK.

The existence of a uropatagium in pterosaurs has been strongly doubted^{3-6,9}, but such a structure is clearly evident in *Sordes* (Fig. 1a, b). It attaches to the inner side of the leg as far as the ankle and is supported along its rear margin by the elongate clawless fifth toe (Fig. 1c, d). The tail is displaced to the left, but the membrane beneath it is undisturbed, confirming Sharov's observation¹¹ that the uropatagium lay below and free of the tail. Fine fibres, extensively distributed within the cheiopatagium, are also present in the uropatagium (Fig. 2a), demonstrating that it too is a flight membrane and not, as some have suggested⁶, a displaced sheet of skin.

What appears to be a small uropatagium is preserved in some examples of *Pterodactylus*^{9,20,21} but direct evidence of this flight surface seems to be lacking in other pterosaurs. However, in many rhamphorhynchoids and early pterodactyloids the hindlimbs of articulated specimens often show a comparable positioning to that seen in *Sordes*: the femur is almost perpendicular

to the spinal column and pronated forward 90°, the anterior surface of the crus faces laterally, and the fifth toe is reflected medially (Fig. 3). This is a most unusual attitude for a tetrapod hindlimb and results from its association with a uropatagium which, with the possible exception of derived pterodactyloids, appears to have been present in many pterosaurs.

The close correlation between skeletal morphology and development of the flight patagia in *Sordes* (Fig. 3) means that reconstructions of the wings in other pterosaurs can now be attempted, even where evidence of soft tissues is not preserved. Preservation of fine internal detail in the flight membranes of *Sordes* also shows that the wings of this and, presumably, other pterosaurs were structurally non-homogenous. The outer half of the wing, filled by long, straight, closely packed fibres (Fig. 2c) seems to have been stiff and relatively inelastic, whereas adjacent to the body the fibres are shorter, more sinuous and loosely packed (Fig. 2a), indicating that the propatagium, uropatagium and

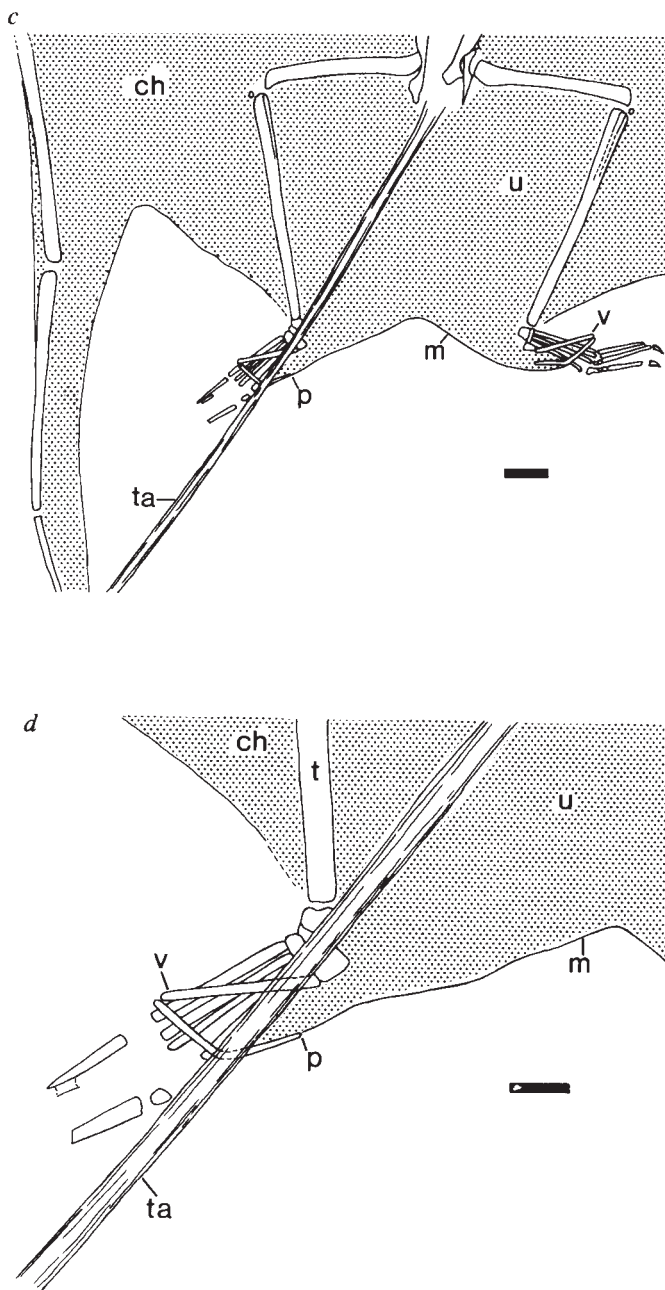
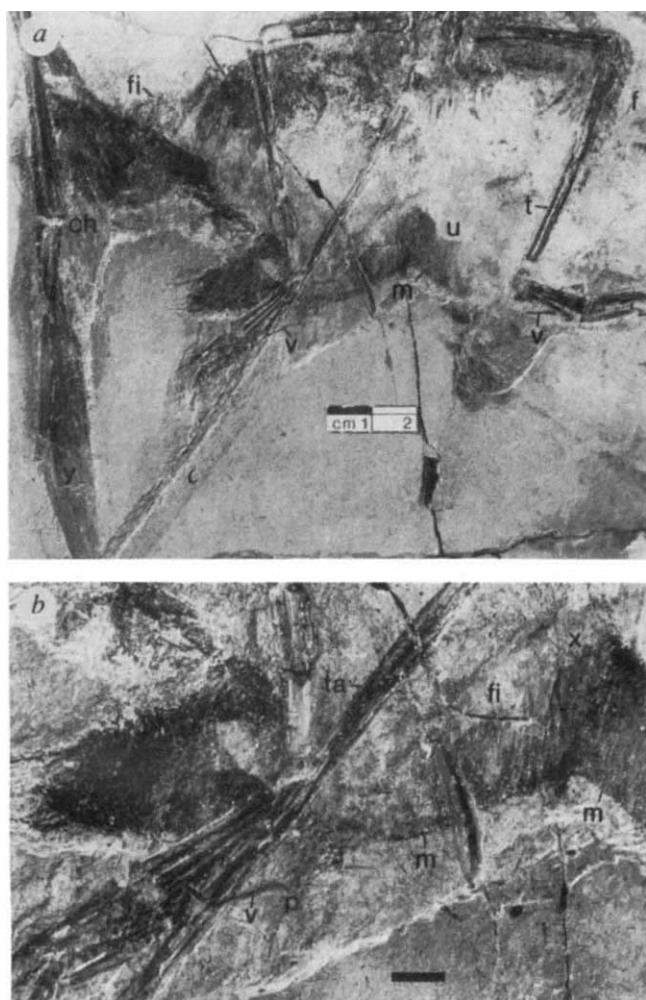


FIG. 1 *Sordes pilosus* (PIN 2585/3). a, Dorsal view of the left wing-finger, the hindlimbs and associated wing membranes. b, Details of the left foot and uropatagium. c and d, Interpretive drawings of a and b (for clarity, only wing membranes and skeletal remains are indicated). The hindlimb is preserved in 'flight' position, with the fifth metatarsal located dorsomedial to the foot, the first phalange of the fifth toe directed laterally, and the second phalange reflected medially to insert into the rear edge of the uropatagium. Areas x and y are shown in detail in Fig. 2. Abbreviations: ch, cheiopatagium; f, fold; fi, structural fibres; m, rear or free margin of uropatagium; p, point of insertion of fifth toe into uropatagium; t, tibia; ta, tail; u, uropatagium; v, fifth toe. Scale bar in b and d, 5 mm; in c, 10 mm.

proximal regions of the cheiropatagium were somewhat softer and more elastic. This model thus provides a neat explanation for the frequently observed absence of evidence of the inner parts of the wings in, for example, Solnhofen pterosaurs^{1,2,6}: only the tough, outer parts were preserved, whereas the softer, inner regions of the wing decayed more rapidly or were not resilient enough to leave recognizable impressions.

Contrary to recent reconstructions in which they are free of the wings¹⁻⁶, the hind limbs were an important component of the flight apparatus (Fig. 3). The configuration and orientation of the cheiropatagium was controlled by the forelimb, assisted by movements of the femur and crus which controlled the position, shape and tension of its relatively elastic proximal region¹⁶. The uropatagium, probably used for manoeuvring and braking,

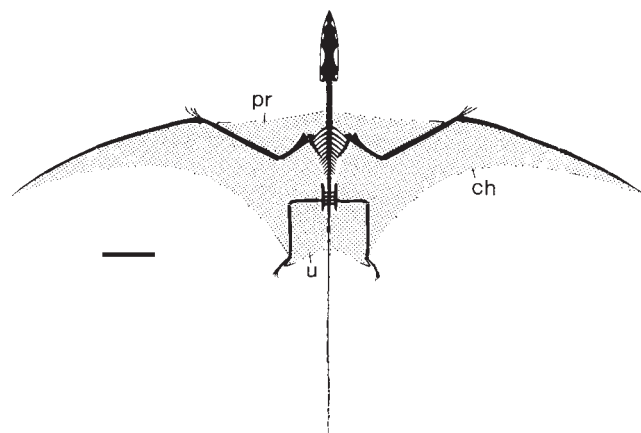


FIG. 3 Restoration of *Sordes pilosus* in dorsal view showing the relationship of the skeleton to the flight membranes. Abbreviations as for Fig. 1; pr, propatagium. Scale bar, 50 mm.

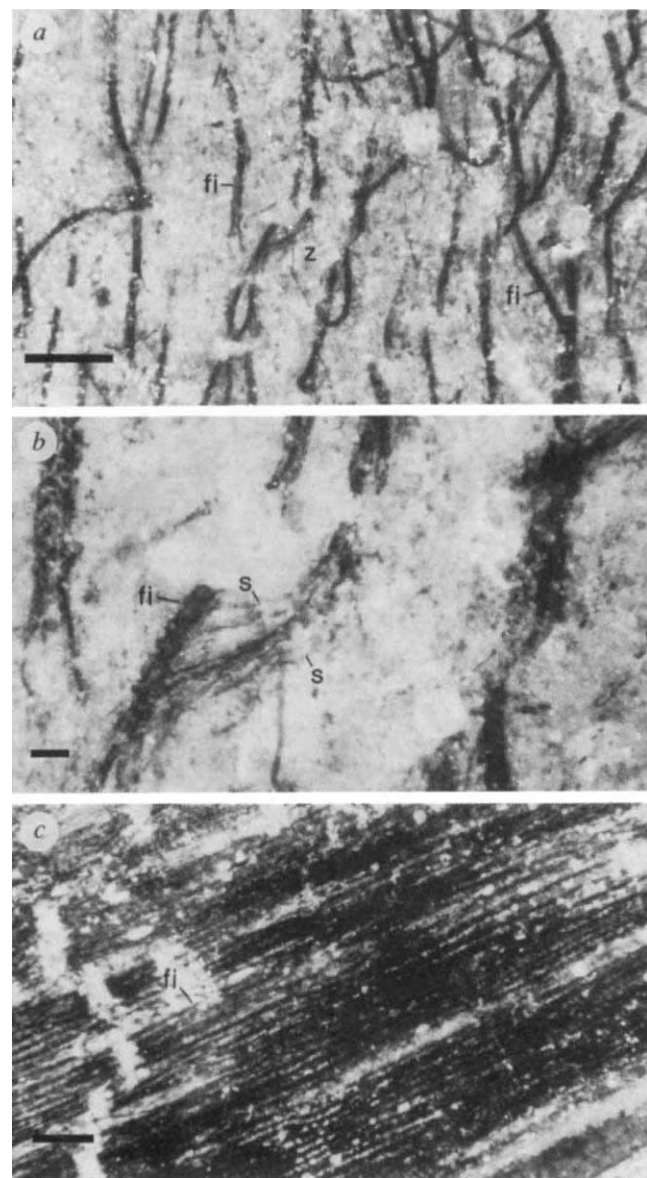


FIG. 2 Details of wing preservation in *Sordes pilosus* (PIN 2585/3). a, Structural fibres in the uropatagium, from a position corresponding to x in Fig. 1b. b, Details of the fibres from position z in a. Each fibre is a composite structure, contrary to recent descriptions of them as solid elements^{5,6}, consisting of numerous ultrafine strands and preserved here as they became partly unravelled. c, Structural fibres in the wing tip, from position y in Fig. 1a. Abbreviations: fi, structural fibres; s, strand. Scale bar in a and c, 1 mm; in b, 0.1 mm.

was manipulated by the fifth toe and thus functioned semi-independently of the cheiropatagium. The highly manipulable flight surface and low wing-loading^{22,23} points to relatively slow, manoeuvrable flight in pterosaurs and is consistent with their probable modes of life as surface feeding piscivores and aerial insectivores^{10,24}. When grounded, the attachment of patagia to the legs and feet must have severely impeded movement. This, and other evidence for a semi-erect, quadrupedal stance and gait^{24,27}, suggests a poor terrestrial ability and a 'gravity-assisted' rather than 'ground up' origin of flight for pterosaurs, particularly as the latter requires the ability to run at relatively high speed²⁸.

The wings of pterosaurs differ in many fundamental ways from those of modern fliers, thus, despite their popularity, models based on birds and bats have failed to provide lasting insights into the construction of the pterosaur flight apparatus. Cases of exceptional preservation such as *Sordes* and others^{1,2,6,9,16,21}, provide a reliable basis for detailed and accurate reconstruction of pterosaur morphology without the need for recourse to arbitrary and potentially misleading analogies⁵. □

Received 8 June; accepted 21 July 1994.

- Wellnhofer, P. *Abh. Bayer. Akad. Wiss. N. F.* **141**, 1–133 (1970).
- Wellnhofer, P. *Palaeontographica* **149**, 1–30 (1975).
- Padian, K. *Paleobiology* **9**, 218–239 (1983).
- Padian, K. *Nat. Hist.* **97**, 58–65 (1988).
- Padian, K. in *Biomechanics in Evolution Sem. Ser. Soc. exp. Biol.* **36** (eds Rayner, J. M. V. & Wootton, R. J.) 146–160 (Cambridge Univ. Press, UK, 1991).
- Padian, K. & Rayner, J. M. V. *Am. J. Sci.* **293A**, 91–166 (1993).
- Pennycuik, C. J. *Mem. Calif. Acad. Sci.* **8**, 83–98 (1986).
- Pennycuik, C. J. *Biol. Rev.* **63**, 209–231 (1988).
- Wellnhofer, P. *Annl. Nat. Hist. Mus. Wien* **88A**, 149–162 (1987).
- Wellnhofer, P. *The Illustrated Encyclopedia of Pterosaurs* (Salamander, London, 1991).
- Sharov, A. G. *Akad. nauk SSSR Paleont. Inst. Tr.* **130**, 104–113 (1971).
- Gekker, R. F. *Akad. nauk SSSR Paleont. Inst. Tr.* **15**, 7–85 (1948).
- Doludenko, M. P., Sakulina, G. V. & Ponomarenko, A. G. *The Geology and Late Jurassic Fauna and Flora of a Unique Locality, Aulie (Karatau Range, Southern Kazakhstan)* (Inst. Geol., Moscow, 1990).
- Riabinkin, A. N. *Akad. nauk SSSR Paleont. Inst. Tr.* **15**, 86–93 (1948).
- Bakhrina, N. N. *Bull. Mos. Soc. nat. Hist. (Geol.)* **63**, 132 (1988).
- Martill, D. M. & Unwin, D. M. *Nature* **340**, 138–140 (1989).
- von Meyer, H. *Zur Fauna der Vorwelt, Vierte Abteilung* (Keller, Frankfurt-am-Main, 1859).
- Broili, F. *Sitzungsber. Bayer. Akad. Wiss., math.-naturwiss.* 23–34 (1925).
- Hazlehurst, G. A. thesis, Bristol Univ. (1991).
- Broili, F. *Sitzungsber. Bayer. Akad. Wiss., math.-naturwiss.* 139–154 (1938).
- Abel, O. *Amer. Mus. Novit.* **192**, 1–12 (1925).
- Bramwell, C. D. & Whitfield, G. R. *Phil. Trans. R. Soc.* **267**, 503–581 (1974).
- Hazlehurst, G. A. & Rayner, J. M. V. *Paleobiology* **18**, 447–463 (1992).
- Wellnhofer, P. & Vahldiek, B. W. *Palaont. Zeitschrift* **60**, 329–340 (1986).
- Unwin, D. M. *Nature* **327**, 13–14 (1987).
- Wellnhofer, P. *Hist. Biol.* **1**, 3–16 (1988).
- Wellnhofer, P. *Palaeontographica* **215**, 43–101 (1991).
- Rayner, J. M. V. in *The Beginnings of Birds* (eds Hecht, M.K., Ostrom, J. H., Viohl, G. & Wellnhofer, P.) 289–292 (Freunde des Jura-Museums Eichstätt, Willibaldsburg, 1985).

ACKNOWLEDGEMENTS. This research was supported by grants from the NERC, the Royal Society and the British Council. We thank A. C. Milner, S. Chapman, C. A. Walker, G. Viohl and P. Wellnhofer for access to specimens; S. C. Bennett, D. Frey, G. Hazlehurst, D. M. Martill and D. Peters for discussion; M. J. Benton, D. E. G. Briggs, E. Buffetaut, K. Padian, J. M. V. Rayner, G. W. Storrs and P. Wellnhofer for critically reading the manuscript; and S. Powell for photography.

Habitat destruction and the extinction debt

David Tilman*, Robert M. May†, Clarence L. Lehman* & Martin A. Nowak†

* Department of Ecology, Evolution and Behavior, 1987 Upper Buford Circle, University of Minnesota, St Paul, Minnesota 55108, USA

† Department of Zoology, Oxford University, South Parks Road, Oxford OX1 3PS, UK

HABITAT destruction is the major cause of species extinctions¹⁻³. Dominant species often are considered to be free of this threat because they are abundant in the undisturbed fragments that remain after destruction. Here we describe a model that explains multispecies coexistence in patchy habitats⁴ and which predicts that their abundance may be fleeting. Even moderate habitat destruction is predicted to cause time-delayed but deterministic extinction of the dominant competitor in remnant patches. Further species are predicted to become extinct, in order from the best to the poorest competitors, as habitat destruction increases. Moreover, the more fragmented a habitat already is, the greater is the number of extinctions caused by added destruction. Because such extinctions occur generations after fragmentation, they represent a debt—a future ecological cost of current habitat destruction.

The predictions are made by an extension of metapopulation models⁵⁻⁹ to multispecies competition⁴. Such models assume that organisms interact within a site, and that sites are linked via dispersal. If species that are competitively superior within sites are poorer dispersers, multispecies coexistence occurs⁴. Such processes seem to explain biodiversity in many animal¹⁰⁻¹⁶ and plant^{4,17-19} communities. The models can be extended to incorporate habitat destruction²⁰. Our model⁴ includes the proportion of sites occupied by species i (p_i), species-specific colonization rates (c_i), mortality (or local extinction) rates (m_i), and habitat destruction (D ; proportion of sites permanently

destroyed). Sites may be as large as local populations^{5-8,12-14} or as small as single individuals⁴. Colonization and loss rates are determined relative to site size. Species are ranked from the best competitor for a limiting resource (species 1) to the poorest. In lieu of competition coefficients, we assume that any occupied site colonized by a superior competitor instantaneously loses the inferior competitor. Thus species i can colonize undestroyed sites not occupied by superior competitors (the first term of equation (1)), but is extinguished from any site invaded by superior competitors (the last term). The equation for species i is:

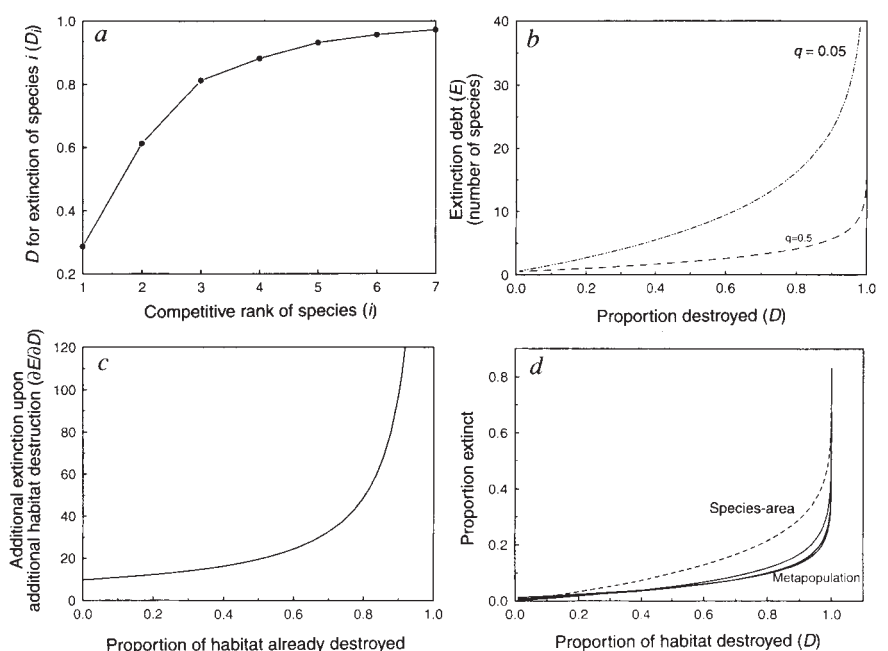
$$\frac{dp_i}{dt} = c_i p_i \left(1 - D - \sum_{j=1}^i p_j \right) - m_i p_i - \sum_{j=1}^{i-1} c_j p_j p_i \quad (1)$$

Its equilibrial ($dp_i/dt=0$) abundance is:

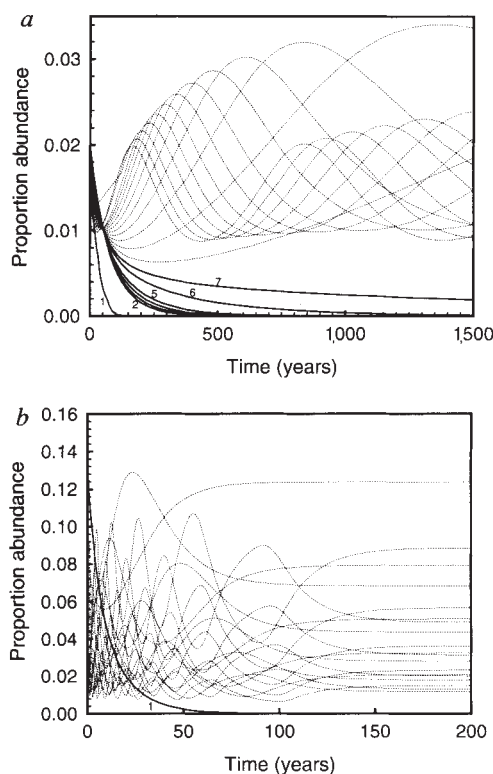
$$\hat{p}_i = 1 - D - \frac{m_i}{c_i} - \sum_{j=1}^{i-1} \hat{p}_j \left(1 + \frac{c_j}{c_i} \right) \quad (2)$$

This must be solved from species 1 to i , with $c_i > 0$ and $\hat{p}_i \geq 0$ for all i . The best competitor has $\hat{p}_1 = 1 - D - m_1/c_1$. It undergoes deterministic extinction once a portion of habitat equal to its abundance in a virgin habitat is destroyed²⁰ because $\hat{p}_1 = 0$ when $D \geq 1 - m_1/c_1$, and $\hat{p}_1 = 1 - m_1/c_1$ when $D = 0$. If the dominant competitor occupies 10% of a virgin habitat, and this proportion of sites is randomly chosen and destroyed, its occupancy of 10% of sites in the undisturbed 90% might seem to assure its survival. Surprisingly, the destruction of random sites has the same eventual effect as selectively destroying precisely those sites occupied by the dominant competitor. Although rare species are always subject to stochastic extinction, this shows that the less abundant a dominant competitor is, the less habitat can be destroyed before its populations in undisturbed fragments are subject to deterministic extinction.

FIG. 1 *a*, Results of an explicitly spatial simulation model of competition in which individuals dispersed locally and interacted only in their site. The habitat was divided into a hexagonal grid of 50×49 sites. Propagules produced at a site dispersed randomly across the four neighbouring rings of sites. In hundreds of such simulations, permanent destruction of a proportion, D , of sites always led first to extinction of the best competitor, followed by the next best competitor, etc., at progressively greater D . Less habitat destruction was required for extinction in simulations than in equations (1–3), apparently because of a percolation effect. *b*, The dependence of E on D , as predicted by equations (1–3). Here, q is the abundance of the best competitor (species 1), and species abundances form a geometric distribution. More species are driven extinct when the dominant competitors are rare ($q=0.05$) than abundant ($q=0.5$), even though total species richness is identical. *c*, The more habitat that is already destroyed, the greater is the effect of additional destruction on extinction. This is based on $\partial E / \partial D = 1 / (2(D-1) \ln(1-q))$. Here $q=0.05$, but the curve has similar shape for all $0 < q < 1$. *d*, The species-area curve for mainland habitats^{2,3} is $S = cA^z$ (where S is species richness, c is a constant, A is area, and z is a constant). We reformulated this to predict the proportion, P_s , of species lost because of habitat destruction, $P_s = 1 - (1-D)^z$ (dashed curve, labelled 'species-area', with $z=0.15$). For the metapopulation model, consider two regions—one species-rich and one species-poor—with roughly the same total numbers of individuals, N , and geometric abundance series. If the population size of the rarest



species is v , then $v = Nq(1-q)^{s-1}$. This gives $S = \ln[v(1-q)/Nq] / \ln[1-q]$. The proportion, P_m , of species lost to destruction is thus $P_m = E/S$. Surprisingly, P_m is fairly independent of q but highly dependent on D (three solid curves; for q of 0.05, 0.2, 0.5).



The D that leads to extinction of other species depends on their traits. Let numerous coexisting species obey equation (1), and let species abundances in a virgin habitat form a geometric series²¹, that is, $p_i = q(1-q)^{i-1}$, where q is the abundance of the best competitor. If all species experience the same loss (m), the required colonization rates⁴ are $c_i = m/(1-q)^{2i-1}$. Substituting into equation (2) and solving for D predicts extinction of the i th species if $D_i \geq 1 - (1-q)^{2i-1}$. Because $D_1 < D_2 < D_3 \dots$, species go extinct from the best to the poorest competitors as habitat destruction increases. Similar results were obtained for three other abundance series and for neighbourhood dispersal (Fig. 1a). Although our model is simple, more realistic elaborations that include explicit spatial structure, slower displacement of inferior competitors, and stochastic effects will not change its key results.

These results are surprising because the species initially most abundant in undisturbed habitat fragments can be the most susceptible to eventual extinction. Habitat destruction lowers effective colonization rates of all species, but most greatly impacts species with lower colonization rates, that is better competitors. Because this requires an interspecific trade-off between colonization and competitive abilities, one test of its applicability is to determine where this trade-off occurs.

Species dynamics after a bout of habitat destruction are shown for cases approximating tropical or temperate forests (Fig.

FIG. 2 Species dynamics after habitat destruction. a, In numerical solution of equation 1, destruction of 1/3 of a habitat led to extinction of the 7 best competitors (solid curves). The 13 inferior competitors (dotted curves) persisted. Parameters were chosen to give a geometric abundance series in a virgin habitat, with the best competitor occupying 3% of sites (that is, $q = 0.03$) and with $m = 0.02 \text{ yr}^{-1}$, mimicking a tropical forest. Lower mortality rates would lead to slower extinctions. b, Same as a, except that the best competitor occupied 20% of the virgin habitat ($q = 0.2$), mimicking a temperate forest. Here only the best competitor (solid curve) was driven extinct by habitat destruction, and the remaining 19 species (dotted curves) persisted.

2a, b). Both cases show the 50 to 400 or more year lag between habitat destruction and extinction of dominant competitors. Because extinctions occur generations after fragmentation, they are a 'debt' that comes due in the future. This model thus provides a mechanism for the slow extinctions that have occurred after habitat isolation, such as on islands created by rising sea levels²²⁻²⁴, or after experimental habitat fragmentation²⁵, or deforestation²⁶.

The number of superior competitors driven extinct by habitat destruction (the extinction debt, E) increases sharply with habitat destruction (Fig. 1b):

$$E = \frac{\ln [(1-D)(1-q)]}{2 \ln [1-q]} \quad (3)$$

This comes from solving for i in the inequality for D_i , because species go extinct in order from 1 to i . A slight increase in habitat destruction threatens many more species if a large portion of a habitat already is destroyed (Fig. 1c). For instance, destroying an additional 1% of habitat causes extinction of 8 times more species if 90% versus 20% of a region already is destroyed. Because E is higher when the best competitor has low abundance (for example, 5%, as in a tropical forest, versus 50%, as in a temperate forest; Fig. 1b), increased habitat destruction threatens more superior competitors for tropical than temperate forests, and for larger and rarer vertebrates than for smaller, more abundant species. Finally, the species-area equation for mainlands and our metapopulation model both predict that the fraction of total species threatened with extinction is a steeply rising function of habitat destruction (Fig. 1d), again showing that additional habitat destruction threatens more species in already fragmented habitats.

Although it is well known that habitat destruction causes extinctions, our results warn that an unanticipated effect of habitat destruction may be the selective extinction of the best competitors. These species are often the most efficient users of resources²⁷ and major controllers of ecosystem functions²⁸. Thus, the extinction debt associated with habitat destruction may have an insidious effect on ecosystems²⁹⁻³⁰. Field work is now required to determine the generality of the requisite competition-colonisation trade-off, the applicability of this theory to particular species and ecosystems, and the spatial scales to which it may apply. □

Received 4 May; accepted 21 July 1994.

1. Ehrlich, P. & Ehrlich, A. *Extinction* (Ballantine Books, New York, 1981).
2. Wilson, E. O. *Biodiversity* (National Academy, Washington DC, 1988).
3. Simberloff, D. *Zh. Obshch. Biol.* **45**, 767-778 (1984).
4. Tilman, D. *Ecology* **75**, 2-16 (1994).
5. Levins, R. & Culver, D. *Proc. natn. Acad. Sci. U.S.A.* **68**, 1246-1248 (1971).
6. Horn, H. S. & MacArthur, R. H. *Ecology* **53**, 749-752 (1972).
7. Cohen, D. & Levin, S. A. *Theo. Pop. Bio.* **39**, 63-99 (1991).
8. Hastings, A. *Theo. Pop. Bio.* **18**, 363-373 (1980).
9. Hanski, I. *Ecology* **64**, 493-500 (1983).
10. Levin, S. A. & Paine, R. T. *Proc. natn. Acad. Sci. U.S.A.* **71**, 2744-2747 (1974).
11. Gaines, S. & Roughgarden, J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3707-3711 (1985).
12. Harrison, S., Murphy, D. D. & Ehrlich, P. R. *Am. Nat.* **132**, 360-382 (1988).
13. Hanski, I. in *Living in a Patchy Environment* (eds Shorrocks, B. & Swingland, I. R.) 127-145 (Oxford Univ. Press, UK, 1990).
14. Shorrocks, B. *Biol. J. Linn. Soc.* **43**, 211-220 (1991).
15. Sale, P. F. *The Ecology of Fishes on Coral Reefs* (Academic, New York, 1991).
16. Doherty, P. & Fowler, T. *Science* **263**, 935-939 (1994).
17. Werner, P. A. & Platt, W. J. *Am. Nat.* **110**, 959-971 (1976).

18. Shmida, A. & Ellner, S. *Vegetatio* **58**, 29-55 (1984).
19. Grubb, P. J. in *Community Ecology* (eds Diamond, J. & Case, T.) 207-226 (Harper & Row, New York, 1986).
20. Nee, S. & May, R. M. *J. Anim. Ecol.* **61**, 37-40 (1992).
21. May, R. M. in *Ecology and Evolution of Communities* (eds Cody, M. L. & Diamond, J. M.) 81-120 (Harvard Univ. Press, Cambridge, MA, 1975).
22. Diamond, J. M. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3199-3203 (1972).
23. Terborgh, J. *BioScience* **24**, 715-722 (1974).
24. Case, T. J., Bolger, D. T. & Richman, A. D. in *Conservat. Biology* (eds Fielder, P. L. & Jain, S. K.) 91-125 (Chapman & Hall, New York, 1992).
25. Lovejoy, T. E. et al. in *Extinctions* (ed. Nitecki, M. H.) 295-325 (Univ. of Chicago Press, Chicago, 1984).
26. Bucher, E. H. *Curr. Ornithol.* **9**, 1-36 (1992).
27. Chapin, F. S. A. *Rev. Ecol. System.* **11**, 233-260 (1980).
28. Pastor, J., Aber, J. D., McCaugherty, C. A. & Melillo, J. M. *Am. Mid. Nat.* **108**, 266-277 (1982).
29. Tilman, D. & Downing, J. A. *Nature* **367**, 363-365 (1994).
30. Naem, S., Thompson, L., Lawler, S., Lawton, J. H. & Woodfin, R. *Nature* **368**, 734-737 (1994).

Positive selection of CD4⁺ T cells by TCR ligation without aggregation even in the absence of MHC

Yousuke Takahama*, Harumi Suzuki†, Kenneth S. Katz†, Michael J. Grusby‡ & Alfred Singer†§

* Syntex Institute of Immunology, Niihari, Japan 300-41

† Experimental Immunology Branch, Building 10, Room 4B-17, National Cancer Institute, Bethesda, Maryland 20892, USA

‡ Department of Cancer Biology, Harvard School of Public Health, and Department of Rheumatology, Harvard Medical School, Boston, Massachusetts 02115, USA

§ To whom correspondence should be addressed

THE developmental fate of immature thymocytes is determined by the specificity of their T-cell antigen receptors (TCRs). Immature CD4⁺8⁺ thymocytes are positively selected to differentiate into mature T cells¹⁻⁶ by recognition of peptides associated with major histocompatibility complex (MHC) encoded molecules⁷⁻¹⁰ on thymic epithelial cells¹¹⁻¹⁴. But neither the identity of molecules transducing positive selection signals nor the nature of the signals themselves is fully known. Here we report that direct ligation of TCR molecules by monoclonal antibodies specific for either clonotypic or CD3 chains can signal immature thymocytes to differentiate into mature CD4⁺8⁺ T cells, even in the absence of MHC expression and MHC-dependent CD4 co-receptor signalling. Moreover, we show that TCR engagement induces positive selection signals only in the absence of TCR aggregation and that TCR aggregation is inhibitory for positive selection. Thus, low valency of TCR crosslinking is a critical parameter¹⁵, distinguishing positive selection from other TCR-mediated signalling events.

To evaluate the role of TCR signalling in positive selection of CD4⁺ T cells, we used mice that were made deficient in MHC class II expression (class II⁻) and so do not generate CD4⁺ T cells^{16,17}. We found that anti-TCR β antibody treatment of fetal thymic organ cultures (FTOC) from class II⁻ mice induced

generation of a distinct population of CD4⁺8⁺ thymocytes (Fig. 1a) and increased significantly the absolute number of CD4⁺8⁺ cells (Table 1). These findings extend observations that CD4⁺8⁺ T cells can be generated in class II⁻ mice by hybrid antibodies that promote binding of TCR⁺ thymocytes to cortical epithelium¹⁸. In addition, we found that anti-TCR β treatment induced CD4⁺8⁺ thymocytes in FTOC from MHC⁻ mice (Fig. 1b) that were deficient in both MHC class I and class II expression^{19,20}. Induction of CD4⁺8⁺ T cells was not limited to antibodies specific for clonotypic TCR chains as a monoclonal antibody specific for an epitope created by pairing of CD3 $\gamma\epsilon$ chains (7D6)²¹ also induced CD4⁺8⁺ T cells (Fig. 1c). The inductive effects of anti-TCR antibodies on CD4⁺8⁺ T cells were in contrast to their overall inhibitory effects in class II⁻ FTOC (Table 1, experiments 1, 2). Reduced recovery of CD4⁺8⁺ thymocytes by anti-TCR β treatment was not unexpected: anti-TCR β antibody inhibits early thymocyte development before the CD4⁺8⁺ stage^{22,23}, resulting in accumulation of CD4⁺8⁺ precursors (Fig. 1); and anti-TCR β antibody induces some CD4⁺8⁺ thymocytes to apoptose^{24,27}. However, our data suggest that anti-TCR β antibody also depletes CD4⁺8⁺ thymocyte number by inducing their further differentiation into CD4⁺8⁺ T cells.

To assess the maturity of induced CD4⁺8⁺ T cells, we examined markers that change during differentiation of CD4⁺8⁺ thymocytes into CD4⁺8⁺ T cells (Fig. 2a). CD5 and CD44 levels were similarly increased on CD4⁺8⁺ T cells relative to CD4⁺8⁺ thymocytes from both normal B6 and antibody-treated class II⁻ FTOC, and levels of thymic shared antigen (TSA-1) were decreased (Fig. 2a). But the most definitive indicator of thymocyte maturity is increased surface-TCR expression. Because antibodies cause internalization of surface TCR complexes, quantification of surface TCR levels after anti-TCR induction required that we remove the antibody one day before harvest to permit thymocytes to re-express surface TCR. We found that CD4⁺8⁺ T cells from normal and class II⁻ FTOCs that had been treated with anti-TCR β antibody expressed identical surface TCR levels which were approximately 10-fold greater than those on their CD4⁺8⁺ precursors (Fig. 2b). Thus, the cell-surface phenotype of antibody induced CD4⁺8⁺ T cells resembled that of CD4⁺8⁺ T cells from normal mice. Even though class II⁻ FTOC were devoid of mature CD4⁺8⁺TCR^{hi} T cells (Fig. 2b,

TABLE 1 Effect of anti-TCR monoclonal antibodies in class II⁻ FTOC*

Expt	Class II ⁻ FTOC†	Specificity of mAb	Thymocyte subset	No. of viable cells recovered per lobe ($\times 10^{-3}$)		
				Medium	mAb-treated	% Control‡
1	16/7d	TCR β	Unfractionated	130	58	45
			CD4 ⁺ 8 ⁺	23	33	143
			CD4 ⁺ 8 ⁺	100	6	6
			CD4 ⁺ 8 ⁺	1.2	9.7	808
			CD4 ⁺ 8 ⁺	6.0	9.4	157
2	16/6d	TCR β	Unfractionated	280	190	68
			CD4 ⁺ 8 ⁺	48	74	154
			CD4 ⁺ 8 ⁺	193	36	18
			CD4 ⁺ 8 ⁺	2.2	32	1,455
			CD4 ⁺ 8 ⁺	36	48	133
3	16/5d	TCR β CD3 $\gamma\epsilon$ CD3 ϵ TCR β + CD3 $\gamma\epsilon$ TCR β + CD3 ϵ	CD4 ⁺ 8 ⁺	0.9	6.8	755
					6.9	766
					0.7	77
					0.4	44
					0.5	55

* Day-16 fetal thymus lobes from class II⁻ mice were placed in organ culture for varying times in the presence or absence of 10–20 $\mu\text{g ml}^{-1}$ of each of the indicated mAbs. Numbers of viable cells recovered per lobe were determined by trypan blue exclusion. Recovered cells were assessed for CD4 and CD8 expression by immunofluorescence and two-colour flow cytometry and the number of recovered cells in each thymocyte subset determined.

† FTOC are indicated as fetal thymic age at setup/number of days of culture, so that 16/7d indicates day-16 fetal thymuses placed in organ culture for 7 days.

‡ Per cent control is $100 \times (\text{number of recovered cells in the presence of mAb}) / (\text{number of recovered cells in the absence of mAb})$.

upper left), we considered that anti-TCR β antibody might function by simply stimulating a few mature CD4⁺8⁺ thymocytes to proliferate. However, newly arising CD4⁺8⁺ thymocytes from anti-TCR β treated class II⁻ FTOC were identical to those from normal B6 FTOC in being negative for the CD25 activation marker, whereas stimulated CD4⁺8⁺ thymocytes were uniformly CD25^{hi} (data not shown). Thus, anti-TCR β antibody does not stimulate proliferation of already mature CD4⁺8⁺ thymocytes in class II⁻ FTOC, but rather generates differentiative signals in immature thymocytes.

To understand how positive selection signals might be distinguished from other TCR signals, we considered that inductive antibodies against TCR β and CD3 $\gamma\epsilon$ each bound to epitopes represented only once in every TCR complex²⁸ and so could directly induce no more than bivalent crosslinking. Consequently, we assessed the effect in FTOC of anti-CD3 ϵ antibody (145-2C11)²⁹ which binds to two epitopes present in every TCR complex^{28,30,31} and so can directly cause multivalent TCR crosslinking. Treatment of class II⁻ FTOC with anti-CD3 ϵ antibody not only failed to generate CD4⁺8⁺ T cells, but it actively inhibited their generation by anti-TCR β antibody (Fig. 3a, lower right). We reasoned that if anti-CD3 ϵ antibody inhibited positive selection because it induced multivalent TCR crosslink-

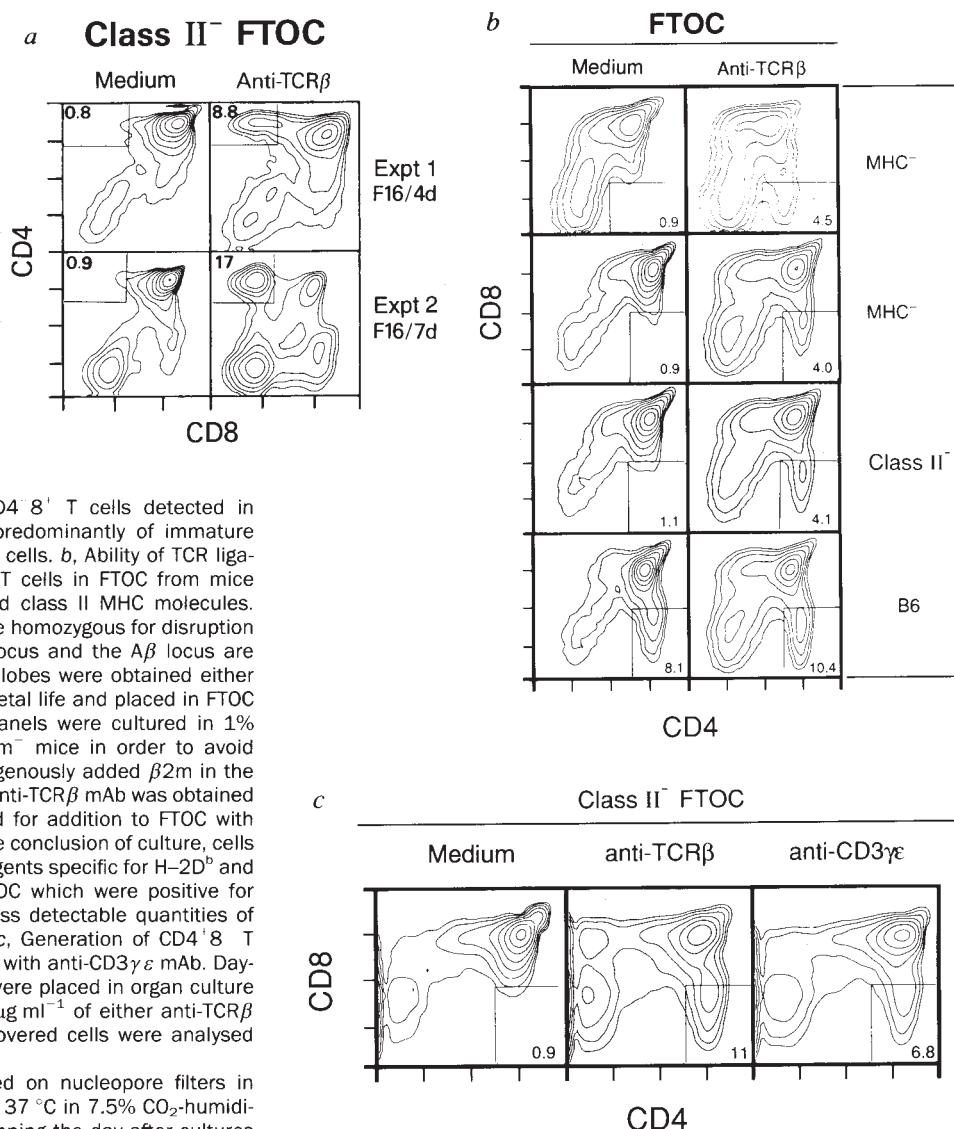
ing, then simultaneous addition of the two inductive antibodies against TCR β and CD3 $\gamma\epsilon$ should also be inhibitory because together they would induce multivalent TCR crosslinking. Indeed, CD4⁺8⁺ T cells were not generated in class II⁻ FTOC treated simultaneously with antibodies against both TCR β and CD3 $\gamma\epsilon$ (Fig. 3a and Table 1, experiment 3).

These experiments indicated that generation of CD4⁺8⁺ thymocytes from class II⁻ FTOC required TCR engagement without extensive crosslinking. However, TCR aggregation can inhibit early thymocyte development before the CD4⁺8⁺ stage^{22,23}. To avoid these early inhibitory effects, thymuses from older neonatal class II⁻ mice whose thymocytes were already at the CD4⁺8⁺ stage were placed into neonatal thymic organ cultures (NTOC) (Fig. 3b). To determine whether multivalent TCR crosslinking would affect induction of CD4⁺8⁺ T cells, we added protein A to crosslink the inductive anti-TCR β antibody. Protein A by itself had no effect on FTOC from either class II⁻ mice (Fig. 3b) or normal B6 mice (not shown). However, protein A inhibited the ability of anti-TCR β antibody to induce CD4⁺8⁺ T cells (Fig. 3b), confirming that multivalent TCR crosslinking ablates an otherwise inductive TCR-positive selection signal.

The present study demonstrates that TCR signalling in immature thymocytes can induce positive selection of CD4⁺8⁺ T cells.

FIG. 1 Generation of CD4⁺8⁺ T cells in thymuses deficient in expression of MHC molecules in response to mAb engagement of surface TCR/CD3 complexes. **a**, Generation of CD4⁺8⁺ T cells in FTOC of class II⁻ mice by treatment with anti-TCR β mAb. Day-16 fetal thymus lobes from class II⁻ H-2^b mice¹⁷ were placed in organ culture for 4 days (experiment 1) or 7 days (experiment 2) with or without 10 $\mu\text{g ml}^{-1}$ anti-TCR β mAb (H57-597)⁴⁴. At the conclusion of the experiment recovered cells were analysed for expression of CD4 and CD8 by two-colour immunofluorescence and flow cytometry. Numbers in each box indicate frequency of cells in that box. It should be noted that class II⁻ mice do contain CD4⁺8⁺ mature T cells but, because they appear late in ontogeny, such cells may be absent from early fetal thymic organ cultures (FTOC). Indeed, CD4⁺8⁺ T cells detected in untreated and mAb-treated FTOCs were predominantly of immature phenotype and were precursors to CD4⁺8⁺ cells. **b**, Ability of TCR ligation to induce the generation of CD4⁺8⁺ T cells in FTOC from mice deficient in expression of both class I and class II MHC molecules. Fetuses derived from matings between mice homozygous for disruption of both the $\beta_2\text{-microglobulin}$ ($\beta_2\text{m}$) gene locus and the A β locus are referred to as MHC⁻ mice²⁰. Fetal thymus lobes were obtained either on day 15 (row 1) or day 16 (rows 2-4) of fetal life and placed in FTOC for 6 days. All FTOC displayed in these panels were cultured in 1% $\beta_2\text{m}^-$ serum that was obtained from $\beta_2\text{m}^-$ mice in order to avoid induction of MHC class I molecules by exogenously added $\beta_2\text{m}$ in the organ cultures. In addition, affinity-purified anti-TCR β mAb was obtained and stored without serum and was diluted for addition to FTOC with medium containing only $\beta_2\text{m}^-$ serum. At the conclusion of culture, cells recovered from FTOC were stained with reagents specific for H-2D^b and I-A^b and, unlike cells from normal B6 FTOC which were positive for both, cells from MHC⁻ FTOC did not express detectable quantities of either MHC class I or class II molecules. **c**, Generation of CD4⁺8⁺ T cells in FTOC of class II⁻ mice by treatment with anti-CD3 $\gamma\epsilon$ mAb. Day-16 fetal thymus lobes from class II⁻ mice were placed in organ culture for 5 days with medium alone or with 10 $\mu\text{g ml}^{-1}$ of either anti-TCR β mAb or anti-CD3 $\gamma\epsilon$ mAb (7D6)²¹, and recovered cells were analysed for expression of CD4 and CD8.

METHODS. Fetal thymic lobes were placed on nucleopore filters in 0.5 ml complete medium and incubated at 37 °C in 7.5% CO₂-humidified air atmosphere as described^{45,46}. Beginning the day after cultures were established, 10 μg affinity-purified mAbs were added daily to the 0.5 ml cultures. The following mAbs were used for staining: anti-CD4 mAb (GK1.5), anti-CD8 mAb (53-6-72) and anti-CD5 mAb (53-7-3),



all from Becton-Dickinson; anti-TSA-1 mAb (MTS.35), anti-CD44 mAb (IM7) and anti-TCR β mAb (H57-597) were from Pharmingen.

Unlike our present results, anti-TCR β mAb was originally reported to block generation of mature T cells in normal FTOC^{32,35}, a result that can now be explained by the presence of aggregates in the antibody preparations which induced multivalent TCR crosslinking and ablated positive selection signals. Indeed, recent studies utilizing different preparations of the identical anti-TCR β mAb have found that it augments generation of mature CD4⁺ T cells in normal FTOC³⁶. In addition, the present results confirm that positive selection of CD4⁺ T cells can occur without MHC-dependent CD4 co-receptor signalling^{18,37}. Nevertheless, CD4 may still function in positive selection of CD4⁺ T cells during normal ontogeny by enhancing TCR avidity for physiological ligands^{37,39}.

Finally, our observation that TCR aggregation is inhibitory for TCR-induced positive selection signals provides a potentially powerful insight into mechanisms of positive selection. Current models of positive selection fall into two apparently incompatible categories predicting that positive selection signals result either from: (1) low-avidity TCR interactions with common self peptides^{7,8,40,41}, or (2) high-avidity TCR interactions with a set of unique selecting peptides^{42,43}. In contrast, on the basis of our results we think that TCR engagement without aggregation

is the critical parameter for positive selection, and not TCR avidity. From our perspective, positive selection signals could be induced by either low-avidity TCR interactions with common self peptides, or by high-avidity TCR interactions with unique selecting peptides. Low-avidity TCR interactions with common self peptides expressed in multiple copies on thymic epithelium would induce positive selection so long as the avidity of interaction was sufficiently low that multiple receptors on an individual thymocyte were not simultaneously engaged and crosslinked. High-avidity TCR interactions with unique selecting peptides would also induce positive selection so long as the selecting peptides were expressed in numbers too low to simultaneously engage and crosslink multiple receptors on an individual thymocyte. Indeed, a prediction of our TCR aggregation model

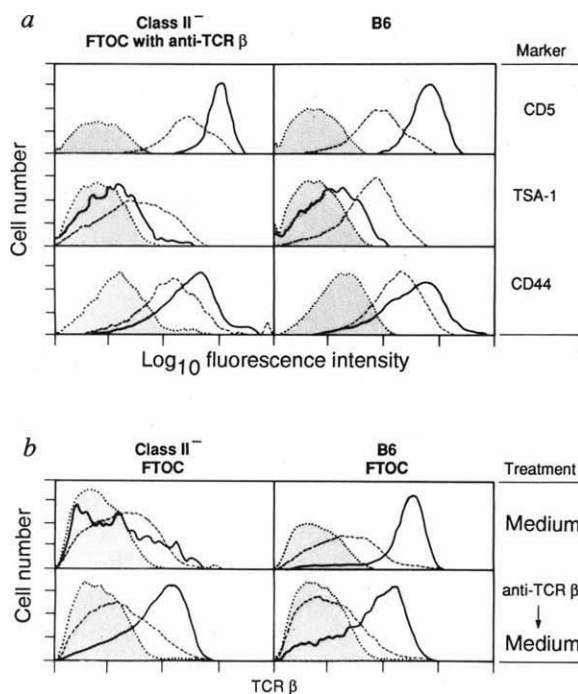


FIG. 2 CD4⁺8⁻ T cells induced in class II⁻ FTOC by anti-TCR β mAb resemble CD4⁺8⁻ T cells generated in normal B6 mice. **a**, Expression of developmentally regulated molecules on CD4⁺8⁻ T cells induced by anti-TCR β mAb. Day-16 fetal thymus lobes from class II⁻ mice were placed in organ culture with 10 μ g ml⁻¹ anti-TCR β mAb, and recovered cells assessed by 3-colour flow cytometry for expression of CD4 and CD8 versus the indicated molecule. Expression of CD5, TSA-1, or CD44 on CD4⁺8⁻ thymocytes (dashed line) and CD4⁺8⁻ T cells (solid line) from mAb-treated class II⁻ FTOC (left panels) and normal B6 thymuses (right panels) are compared with those stained with a negative control mAb (shaded). **b**, Expression of TCR β on CD4⁺8⁻ T cells induced by anti-TCR β mAb. Where indicated, day-16 fetal thymus lobes from either class II⁻ or normal B6 were either cultured for 6 days with medium alone or were cultured for 5 days in the presence of 10 μ g ml⁻¹ anti-TCR β mAb, extensively washed, transferred to filters on fresh culture medium, and incubated for an additional 24 h without mAb to permit re-expression of surface TCR complexes. Recovered cells were assessed by 3-colour flow cytometry for expression of CD4 and CD8 versus TCR β or CD3- ϵ . Expression of TCR β on CD4⁺8⁻ thymocytes (dashed line) and CD4⁺8⁻ T cells (solid line) are compared with those stained with a negative control mAb (shaded). Staining with mAbs against TCR β and CD3- ϵ yielded identical results.

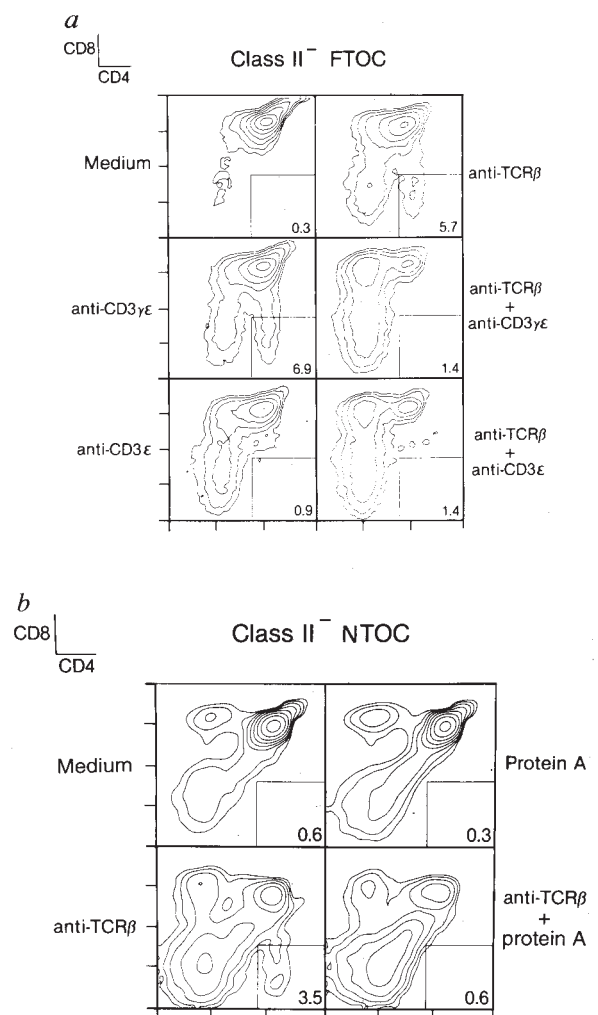


FIG. 3 Induction of CD4⁺8⁻ T cells requires TCR engagement without extensive crosslinking. **a**, Day-16 fetal thymus lobes were cultured for 5 days in medium alone or in the presence of 10 μ g ml⁻¹ of each indicated mAb, and recovered cells were analysed for expression of CD4 and CD8 by two-colour flow cytometry. Numbers in each box indicate frequency of cells in that box. Where it is indicated that mixtures of two mAbs were added to FTOC, the final concentration of each mAb was 10 μ g ml⁻¹. The inhibitory effects of mAb mixtures were not simply due to increased amounts of mAb in the cultures, as individual mAbs were equally inductive at final concentrations of either 10 μ g ml⁻¹ or 20 μ g ml⁻¹ (not shown). **b**, Thymus lobes from class II⁻ mice were obtained on the day of birth and placed in neonatal thymic organ culture (NTOC) for 4 days in medium alone, 10 μ g ml⁻¹ protein A, 10 μ g ml⁻¹ anti-TCR β mAb, or both protein A and anti-TCR β mAb. Recovered cells were analysed for expression of CD4 and CD8 by two-colour flow cytometry. Numbers in each box indicate frequency of cells in that box.

is that negatively selecting self peptides in concentrations too low to cause receptor aggregation would induce positive selection of high-avidity TCR, a prediction that has now been fulfilled^{9,10}. □

Received 11 April; accepted 15 July 1994.

1. Teh, H. S. et al. *Nature* **335**, 229–233 (1988).
2. Sha, W. C. et al. *Nature* **336**, 73–76 (1988).
3. Berg, L. J. et al. *Cell* **58**, 1035–1046 (1989).
4. Pircher, H., Burki, K., Lang, R., Hengartner, H. & Zinkernagel, R. M. *Nature* **342**, 559–561 (1989).
5. Blackman, M., Kappler, J. & Marrack, P. *Science* **248**, 1335–1341 (1990).
6. Kaye, J. & Ellenberger, D. L. *Cell* **71**, 423–435 (1992).
7. Ashton-Rickardt, P. G., van Kaer, L., Schumacher, T. N. M., Pleogh, H. L. & Tonegawa, S. *Cell* **73**, 1041–1049 (1993).
8. Hogquist, K. A. et al. *Cell* **76**, 17–27 (1994).
9. Ashton-Rickardt, P. G. et al. *Cell* **76**, 651–663 (1994).
10. Sebzda, E. et al. *Science* **263**, 1615–1618 (1994).
11. Lo, D. & Sprent, J. *Nature* **319**, 672–675 (1986).
12. Vukmanovic, S., Granda, A. G., Faas, S. J., Knowles, B. B. & Bevan, M. J. *Nature* **359**, 729–732 (1992).
13. Hugo, P., Kappler, J. W., Godfrey, D. I. & Marrack, P. C. *Nature* **360**, 679–682 (1992).
14. Anderson, G., Jenkinson, E. J., Moore, N. C. & Owen, J. J. T. *Nature* **362**, 70–73 (1993).
15. Dintzis, H. M., Dintzis, R. Z. & Vogelstein, B. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3671–3675 (1976).
16. Cosgrove, D. et al. *Cell* **66**, 1051–1066 (1991).
17. Grusby, M. J., Johnson, R. S., Papaioannou, V. E. & Glimcher, L. H. *Science* **253**, 1417–1420 (1991).
18. Muller, K. P. & Kyewski, B. A. *Eur. J. Immun.* **23**, 1661–1670 (1993).
19. Chan, S., Cosgrove, D., Watzinger, C., Benoist, C. & Mathis, D. *Cell* **73**, 225–236 (1993).
20. Grusby, M. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3913–3917 (1993).
21. Coulie, P. G. et al. *Eur. J. Immun.* **21**, 1703–1709 (1991).

22. Hunig, T. *Eur. J. Immun.* **18**, 2089–2092 (1988).
23. Takahama, Y. & Singer, A. *Science* **258**, 1456–1462 (1992).
24. Smith, C. A., Williams, C. T., Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **337**, 181–184 (1989).
25. McConkey, D. J., Hartzell, P., Amador-Perez, J. F., Orrenius, S. & Jondal, M. J. *Immun.* **143**, 1801–1806 (1989).
26. Murphy, K. M., Heimberger, A. B. & Loh, D. Y. *Science* **250**, 1720–1722 (1990).
27. Nakayama, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9949–9953 (1991).
28. Punt, J. A., Roberts, J. L., Kearse, K. P. & Singer, A. J. *exp. Med.* **180**, 587–593 (1994).
29. Leo, O., Foo, M., Sachs, D. H., Samelson, L. E. & Bluestone, J. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1374–1378 (1987).
30. Blumberg, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7220–7224 (1990).
31. de la Hera, A., Muller, U., Olsson, C., Isaacs, S. & Tunnacliffe, A. J. *exp. Med.* **173**, 7–17 (1991).
32. McDuffie, M., Born, W., Marrack, P. & Kappler, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8728–8732 (1986).
33. Born, W. et al. *J. Immun.* **138**, 999–1008 (1987).
34. Finkel, T. H. et al. *Cell* **58**, 1047–1054 (1989).
35. Yachelini, P., Falk, I. & Eichmann, K. J. *Immun.* **145**, 1382–1389 (1990).
36. Punt, J. A., Hosono, M. & Hashimoto, Y. J. *Immun.* **151**, 1290–1302 (1993).
37. Killeen, N. & Littman, D. R. *Nature* **364**, 729–732 (1993).
38. Swain, S. L. *Immun. Rev.* **74**, 129–142 (1983).
39. Doyle, C. & Strominger, J. L. *Nature* **330**, 256–259 (1987).
40. Singer, A., Mizuochi, T., Munitz, T. I. & Gress, R. E. in *Progress in Immunology VI* (eds Cinader, B. & Miller, R. G.) 60–66 (Academic, New York, 1986).
41. Singer, A., Munitz, T. I. & Gress, R. E. *Transplant. Proc.* **19**, 107–110 (1987).
42. Marrack, P., McCormack, J. & Kappler, J. *Nature* **338**, 503–505 (1989).
43. Kourilsky, P. & Clavierie, J. M. *Cell* **56**, 327–329 (1989).
44. Kubo, R. T., Born, W., Kappler, J. W., Marrack, P. & Pigeon, M. J. *Immun.* **142**, 2736–2742 (1989).
45. Robinson, J. H. & Owen, J. J. T. *Clin. Exp. Immun.* **27**, 322–327 (1976).
46. Mandel, T. E. & Kennedy, M. M. *Immunology* **35**, 317–331 (1978).

ACKNOWLEDGEMENTS. We thank P. Kourilsky for discussion; L. Granger, C. Johnson and S. Sharrow for help with flow cytometry; and R. Hodes, P. Kourilsky, J. Punt, S. Sharrow and D. Singer for critically reading the manuscript. Y.T. and H.S. contributed equally to this work.

M and L cones in macaque fovea connect to midget ganglion cells by different numbers of excitatory synapses

David J. Calkins, Stanley J. Schein*, Yoshihiko Tsukamoto† & Peter Sterling

Department of Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

* Department of Psychology, University of California, Los Angeles, California 90024, USA

† Department of Anatomy, Hyogo College of Medicine, Nishinomiya, 663 Hyogo, Japan

VISUAL acuity depends on the fine-grained neural image set by the foveal cone mosaic^{1–3}. To preserve this spatial detail, cones transmit through non-divergent pathways: cone→midget bipolar cell→midget ganglion cell. Adequate gain must be established along each pathway; crosstalk and sources of variation between pathways must be minimized. These requirements raise fundamental questions regarding the synaptic connections: (1) how many synapses from bipolar to ganglion cell transmit a cone signal and with what degree of crosstalk between adjacent pathways; (2) how accurately these connections are reproduced across the mosaic; and (3) whether the midget circuits for middle (M) and long (L) wavelength sensitive cones are the same. We report here that the midget ganglion cell collects without crosstalk either 28 ± 4 or 47 ± 3 midget bipolar synapses. Two cone types are defined by this difference; being about equal in number and distributing randomly in small clusters of like type, they are probably M and L.

The non-divergent circuit from each foveal cone actually involves two midget bipolar cells⁴. Their axons descend in parallel to different levels^{5,6} of the inner plexiform layer where they contact the dendritic arbors of single 'off' and 'on' midget ganglion cells^{7,8}. We reconstructed such paired midget bipolar and ganglion cells from electron micrographs of serial sections (Fig.

1a). Studying 95 off and 91 on midget bipolar axon terminals, we found that each contacted one midget ganglion cell exclusively without contributing even a single synapse to a neighbouring midget ganglion cell. Thus, in macaque fovea the midget bipolar pathways from adjacent cones neither diverge nor converge. This has also been reported for four midget bipolar cells in human parafovea⁹.

We quantified the bipolar synapses to the midget ganglion cells. Each synapse was identified by a ribbon pointing between a pair of postsynaptic processes ('dyad'¹⁰). Most dyads comprised a midget ganglion cell dendrite and an amacrine process (88%); however, some dyads comprised two amacrine processes (9%) or two ganglion cell dendrites (1%). Occasionally, one postsynaptic element of a dyad was a non-midget ganglion cell dendrite (2%). The off and on bipolar terminals connected to the same cone had similar numbers of ribbon synapses. For example, the left pair in Fig. 1b had 25 and 32 ribbons, and the right pair had 46 and 52 ribbons. Figure 1c shows this strong association for 56 such pairs (Spearman rank correlation coefficient¹¹ $r_s = 0.76$, $P < 0.01$).

When we enumerated the ribbon synapses of every complete off and on midget bipolar terminal in the tissue, two non-overlapping distributions emerged: one with 32 ± 3 ribbons and one with 49 ± 3 ribbons (means $\pm$ s.d.: Fig. 1d). The Kruskal 'dip intensity' statistic¹² for this distribution is 4.0, indicating strong bimodality ($P < 0.01$, $n = 145$). The number of bipolar synapses to the ganglion cells differed correspondingly: 28 ± 4 versus 47 ± 3 synapses. Each midget ganglion cell also received amacrine cell synapses about equal in number to the ribbon synapses: 30 ± 3 versus 45 ± 3 synapses. Equal amacrine input was also reported for midget ganglion cells in human parafovea⁹. The bipolar terminals with more synapses had greater membrane surface areas, $103 \pm 10 \mu\text{m}^2$ as opposed to $79 \pm 9 \mu\text{m}^2$, and so did the corresponding ganglion cell dendrites, $105 \pm 23 \mu\text{m}^2$ versus $81 \pm 7 \mu\text{m}^2$. These differences were significant for both the bipolar terminals (Kruskal-Wallis¹¹ $H = 8.31$, $P < 0.005$, d.f. = 1) and the ganglion cells dendrites ($H = 4.33$, $P = 0.04$, d.f. = 1). The off and on bipolar-to-ganglion cell junctions from the same cone had the same membrane surface areas for both bipolar terminals ($H = 0.23$, $P = 0.66$, d.f. = 1) and for ganglion cell dendrites ($H = 0.02$, $P = 0.90$, d.f. = 1; compare ref. 13). Occasionally, a midget ganglion cell received a synapse from a diffuse bipolar cell terminal (see also ref. 9).

In retrospect, Polyak's classic drawings of Golgi-impregnated cells in rhesus monkey and also in chimpanzee suggest two sizes of dendritic arbor (Figs 70, 74 in ref. 7). Also in retrospect, the four midjet ganglion cells reconstructed from human parafovea also had small and large junctions: 55 and 57 versus 71 and 81 bipolar synapses⁹. This fits the hypothesis of two types because variation between cells of the same type is generally about 5–10%^{14–16}. Thus, the two sizes of midjet junction observed here in *Macaca fascicularis* may be general in Old World primates.

Fifty-nine cones in this tissue were associated with small junctions and 45 cones were associated with large junctions, yielding a ratio of 1.3. The ratio in the population represented by the sample lies between 0.9 and 2.1 with 95% confidence.

The mosaic formed by small and large junctions represents the arrangement of cone outer segments at 1° (Fig. 1e). This is so because the axons of adjacent cones do not cross¹⁷, nor do

those of adjacent midjet bipolar cells. The mosaic contains small clusters of like type. Such clusters would be expected from a random distribution. This we verified by measuring, along each axis of the triangular lattice, the number and lengths of unbroken sequences ('strings') of like cone type. The number of strings summed over all three axes was not different from that expected from a random mosaic ($Z = -0.88$, $P = 0.38$)¹¹. The lengths of strings, 2.0 ± 1.3 cones for cones with small midjet junctions and 1.4 ± 0.7 for cones with large junctions, were not different from those calculated for five random mosaics with the same number and ratio of cones (Fig. 2).

The finding of two cone types, about equal in number and random in distribution, matches closely the pattern of M and L cones shown by *in situ* measurements of spectral sensitivity¹⁸. Therefore, we speculate that the two cone types identified here by differences in midjet circuitry are M and L; indeed, we can

FIG. 1 a, Midjet circuit from one cone. Cone terminal (orange) connects via paired midjet bipolar cells (off, cyan; on, violet) to paired midjet ganglion cells (off, pink; on, yellow); the cone also connects to 'diffuse' bipolar neurons not shown. Off and on regions of the inner plexiform layer^{5,6} are demarcated. b, Left, paired midjet pathways for one cone. Bipolar terminals contained 25 and 32 ribbon synapses (white rectangles); corresponding midjet ganglion cells received 22 and 31 bipolar synapses (white circles). Right, midjet pathways from a neighbouring cone. Bipolar terminals contained 46 and 52 ribbons; corresponding midjet ganglion cells received 43 and 51 bipolar synapses. Some synapses are hidden. c, Off and on midjet bipolar terminals connected to the same cone have similar numbers of ribbons. d, Number of midjet bipolar ribbon synapses formed two non-overlapping distributions. e, Inferred mosaic of 108 cones at 1° nasal fovea. S cones, identified by their lack of an on midjet circuit^{19,20} and by connections to on 'blue cone bipolar'^{21,22} cells, are coloured blue. Cones connected to small midjet junctions are red ($N = 59$). Cones connected to large midjet junctions are green ($N = 45$). This mosaic shows the putative distribution of M and L cones, but the colour that actually represents M or L could not be established.

METHODS. A retina was obtained from an adult male *Macaca fascicularis* and prepared for electron microscopy²³. Consecutive sections (319) were cut radially at 90° nm through the nasal fovea along the horizontal meridian and photographed at $2,000\times$ magnification, yielding a patch $30 \times 130 \mu\text{m}$, with midjet circuits at $510\text{--}640 \mu\text{m}$ eccentricity and the cone inner segments that serve them at $0.8\text{--}1.2^\circ$. Neurons were reconstructed by tracing their successive profiles onto mylar sheets. The tracings were stacked by computer²⁴ and surface-rendered (a, b). For 43 midjet pathways 1:1 correspondence between bipolar terminal and ganglion cell was assessed directly by tracing both cells. Fifty-five cones were represented in the series only by their midjet cell pathways. These were quantified as described and projected by interpolation onto the mosaic in e.

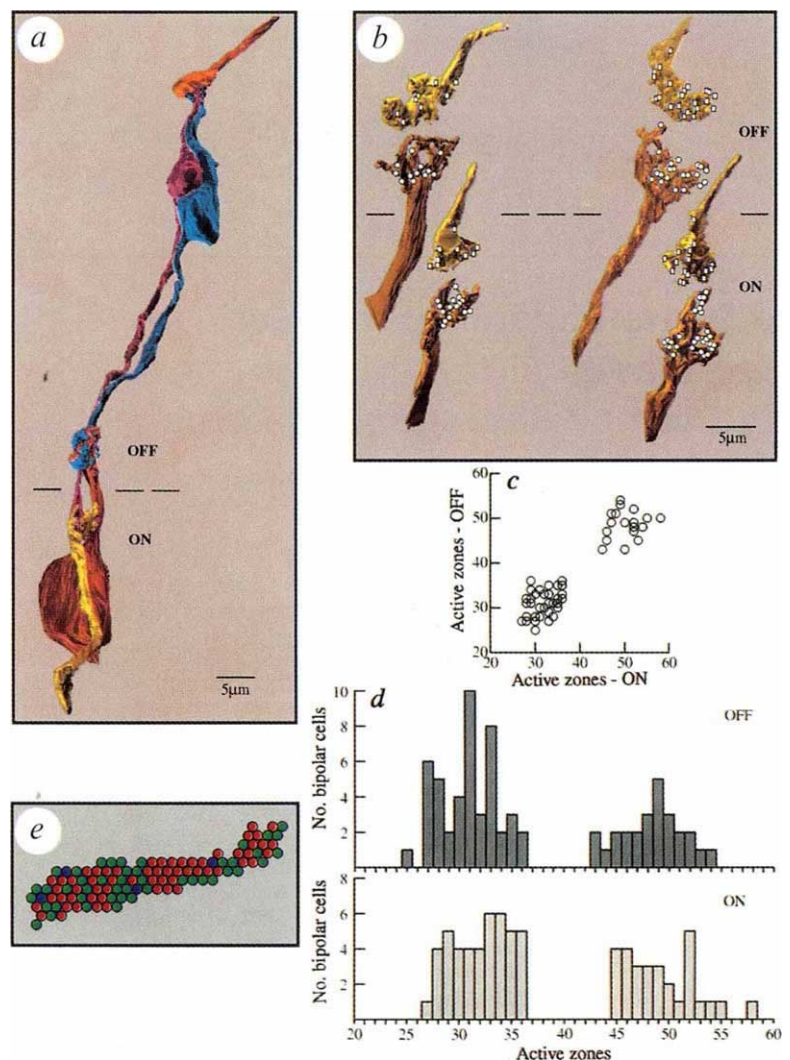
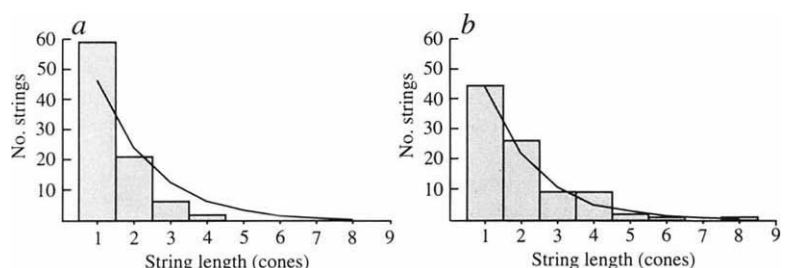


FIG. 2 a and b, Distribution of lengths of unbroken sequences ('strings') of cones connected to large midjet circuits (in a, $N = 88$ strings) and to small midjet circuits (in b, $N = 92$ strings). Smooth curves are calculated from the simple binomial model, Np^i , where N is the number of strings, i is string length and p is the best-fitting probability of success parameter. With N fixed, the sum of squares for error was minimum with $p = 0.52$ (a) and 0.50 (b), suggesting the possibility that the cone sample was drawn from a population with equal numbers of M and L cones. Five random mosaics with the same number and ratio of cones were generated; the distributions of string lengths for the two cone types in the inferred mosaic (Fig. 1e) were not statistically different from those of the random mosaics ($\chi^2 = 0.02$, $P = 0.90$, d.f. = 1).



think of no plausible alternative. Whether M corresponds to the large midget junctions and L to small junctions, or vice versa, we cannot yet say, but because primate ganglion cells can now be identified following *in situ* spectral measurements⁶, this issue can be addressed.

Selective pressure is likely to minimize the number of synapses employed in transferring the neural image from the cone mosaic to the ganglion cells. Therefore, the 28 ± 4 synapses from bipolar to ganglion cell (glutamatergic and therefore excitatory) is probably irreducible. Also, it is likely that there is pressure to minimize the variation between off and on midget pathways leading from the same cone (Fig. 1*b, c*) in order that each point in the neural image be represented equally well for signals dimmer than

or brighter than the mean. For the midget junctions serving cones of the same type, variation is equally small (s.d. is about 10% of the mean). This, too, seems critical in order that neighbouring points in the natural image be represented equally well. Such constancy of synaptic connections for neurons of the same type has been demonstrated for many other cell types in mammalian retina¹⁴⁻¹⁶.

It might also have been expected that the midget circuits leading from M and L cones use the same number of synapses, but apparently one type employs about 50% more excitatory synapses than the other. This difference is proving helpful in determining the spectral characteristics of other retinal circuits, but why it should be so is an enigma. □

Received 19 May; accepted 12 July 1994.

- Williams, D. R. *Trends Neurosci.* **9**, 193-198 (1986).
- DeValois, R. L., Morgan, H. & Snodderly, D. M. *Vis. Res.* **14**, 75-81 (1974).
- Merigan, W. H. & Katz, L. M. *Vis. Res.* **30**, 985-991 (1990).
- Kolb, H. *Phil. Trans. R. Soc. Lond. B* **258**, 261-283 (1970).
- Nelson, R., Famigletti, E. V. & Kolb, H. *J. Neurophysiol.* **41**, 472-483 (1978).
- Dacey, D. & Lee, B. B. *Nature* **367**, 731-735 (1994).
- Polyak, S. L. *The Retina* (Univ. Chicago Press, Chicago, 1941).
- Boycott, B. B. & Dowling, J. E. *Phil. Trans. R. Soc. Lond. B* **255**, 109-184 (1969).
- Kolb, H. & Dekorver, L. *J. comp. Neurol.* **303**, 617-636 (1991).
- Dowling, J. E. & Boycott, B. B. *Proc. R. Soc. Lond. B* **166**, 80-111 (1966).
- Mendenhall, W., Wackerly, D. D. & Scheaffer, R. L. *Mathematical Statistics with Applications* 4th edn 708-715 (Duxbury, Belmont, CA, 1990).
- Giacomelli, F., Wiener, J., Kruskal, J. B., Pomeranz, J. V. & Loud, A. V. *J. Histochem. Cytochem.* **19**, 426-433 (1971).
- Dacey, D. M. & Petersen, M. R. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9666-9670 (1992).

- McGuire, B. A., Stevens, J. K. & Sterling, P. *J. Neurosci.* **4**, 2920-2938 (1984).
- Cohen, E. & Sterling, P. *Eur. J. Neurosci.* **4**, 506-520 (1992).
- Anhelt, P., Keri, C. & Kolb, H. *J. comp. Neurol.* **293**, 39-53 (1990).
- Wässle, H., Boycott, B. B. & Röhrenbeck, J. *Eur. J. Neurosci.* **1**, 421-435 (1989).
- Mollon, J. D. & Bowmaker, J. K. *Nature* **360**, 677-679 (1992).
- Klug, K., Tsukamoto, Y., Sterling, P. & Schein, S. *J. Soc. for Neurosci. Abstr.* 352.5 (Anaheim, CA, 1992).
- Klug, K., Tsukamoto, Y., Sterling, P. & Schein, S. *J. ARVO Abstr.* 1398 (Association for Research in Vision and Ophthalmology, Sarasota, FL, 1993).
- Mariani, A. P. *Nature* **308**, 184-186 (1984).
- Kouyama, N. & Marshak, D. W. *J. Neurosci.* **12**, 1233-1252 (1992).
- Tsukamoto, Y., Masaracha, P., Schein, S. J. & Sterling, P. *Vis. Res.* **32**, 1809-1815 (1992).
- Smith, R. G. *J. Neurosci. Meth.* **21**, 55-69 (1987).

ACKNOWLEDGEMENTS. We thank P. Masaracha for preparing the material for this project, K. Klug for the S-cone terminal identification, A. Shrom for assistance with computer graphics, T. Chou for the surface-rendering algorithm, and W. Ewens for discussion.

p53-dependent apoptosis produced by Rb-deficiency in the developing mouse lens

Sharon D. Morgenbesser*, Bart O. Williams†, Tyler Jacks†† & Ronald A. DePinho*

* Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461, USA
† Howard Hughes Medical Institute, Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

THE retinoblastoma tumour-suppressor gene (*RB*) has been implicated in negative growth regulation, induction of differentiation, and inhibition of cellular transformation¹. Homozygous inactivation of the *Rb* gene in the mouse leads to mid-gestational lethality with defects in erythropoiesis and neurogenesis²⁻⁴. Here we describe the effects of the *Rb*-deficient state on the development of the ocular lens. The regional compartmentalization of growth, differentiation and apoptosis in the developing lens provides an ideal system to examine more closely the relationships of these processes *in vivo*. We demonstrate that loss of *Rb* function is associated with unchecked proliferation, impaired expression of differentiation markers, and inappropriate apoptosis in lens fibre cells. In addition, we show that ectopic apoptosis in *Rb*-deficient lenses is dependent on *p53*, because embryos doubly null for *Rb* and *p53* show a nearly complete suppression of this effect. This developmental system provides a framework for understanding the consequences of the frequent mutation of both *RB* and *p53* in human cancer.

In an effort to establish a biological system to study *Rb* function, we and others have used gene targeting to inactivate *Rb* in the mouse²⁻⁴. Mice heterozygous for the *Rb*^{Δ31} allele (*Rb*^{+/-}) are grossly normal but are highly predisposed to the development of pituitary adenocarcinomas, and inheritance of two mutant *Rb* alleles (*Rb*^{-/-}) results in lethality at 13-15 days of gestation³.

To characterize more precisely the role of *Rb* in the transition from active proliferation to end-stage differentiation as well as to investigate the pathological consequences of alterations in this process, we examined the effects of *Rb* deficiency on ocular lens development⁵. During normal mouse lens development, following lens vesicle formation on embryonic day 11.5 (E11.5), lens cells positioned on the posterior wall of the vesicle withdraw from the cell cycle and elongate in an anterior direction into the vesicle. By E14.5, elongation is complete and results in a lens consisting of postmitotic differentiated fibre cells that are covered anteriorly by a layer of proliferating, immature epithelial cells (Fig. 1*a*). For the purposes of this study, it is important to note that although lens epithelial cells occasionally undergo apoptosis, lens fibre cells do not^{6,7} (Fig. 4*a*).

Histological analysis of *Rb*^{-/-} lenses from E12.5-E14.5 revealed a significant increase in the number of nuclei as well as occasional mitotic figures in the lens fibre cell compartment (Fig. 1*a, b*). That loss of the *Rb* protein is associated with inappropriate progression through S-phase in lens fibre cells was confirmed by the large number of cells positive for 5-bromo-2'-deoxyuridine (BrdU) throughout the lens fibre cell compartment in *Rb*^{-/-} embryos (Fig. 2*b*), but not in wild-type controls (Fig. 2*a*). Given that withdrawal from the cell cycle appears to be a requirement for normal differentiation in many cell types, we examined whether continued proliferation in *Rb*^{-/-} lens fibre cells affected their ability to differentiate terminally. Although loss of *Rb* appeared to have a minimal impact on the gross morphological formation of the lens, the highly organized parallel and extended array of normally differentiated lens fibre cells was altered and replaced by a disordered arrangement in *Rb*^{-/-} lenses (Fig. 1*a, b*).

As normal lens fibre cell development is dependent upon the proper temporal and spatial expression of α , β and γ -crystallins and the major intrinsic protein of the lens fibre plasma membrane, MIP26 (refs 8, 9), immunofluorescence was used to assay the expression of these stage-specific proteins. Lack of *Rb* protein did not appear to affect the regional distribution of α -crystallin, an early-stage marker, although the level of staining was mildly reduced in the *Rb*^{-/-} lenses relative to wild-type controls (Fig. 3*a, b*). In contrast, expression of β - and γ -crystallins and MIP26, the late-stage markers^{8,10}, was markedly

† To whom correspondence should be addressed.

decreased in the $Rb^{-/-}$ lens (Fig. 3, c–h). The more severe reduction in the levels of late-stage markers relative to that of an early-stage marker coupled with the formation of an intact lens structure suggests that the more advanced aspects of differentiation, rather than commitment, are perturbed in the $Rb^{-/-}$ mouse lens.

Previous histological analysis of Rb -deficient embryos revealed widespread cell death in the peripheral and central nervous systems^{2,4}. Confocal microscopic analysis of propidium-iodide-stained wild-type and $Rb^{-/-}$ lens sections verified the presence of hallmark features of apoptosis, including nuclear fragmentation and chromatin condensation along the nuclear envelope, in many $Rb^{-/-}$ lens fibre cells (data not shown). To

establish further that the mechanism of lens fibre cell death was apoptosis, rather than necrosis, we assayed for the presence of excessive DNA 3'-OH ends using the TdT-mediated dUTP-biotin nick-end labelling (TUNEL) method¹¹. As anticipated, there was a complete absence of TUNEL-stained nuclei in wild-type lenses (Fig. 4a). In contrast, nuclear staining in many lens fibre cells of $Rb^{-/-}$ lenses confirmed the presence of apoptosis in this compartment (Fig. 4b) and provided the first clear evidence that the massive cell death that characterizes the Rb -deficient phenotype occurs via this programmed mechanism.

Studies in cultured cells have established that $p53$ is required for apoptosis associated with deregulated growth^{12–15}. However, it is also clear that apoptosis can proceed via $p53$ -dependent

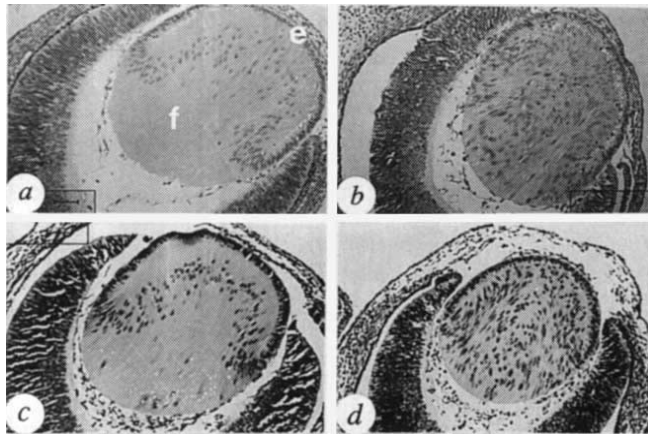


FIG. 1 Morphological development of lenses in E13.5 wild-type and mutant embryos. a, Wild-type; b, $Rb^{-/-}$; c, $p53^{-/-}$; and d, $Rb^{-/-}; p53^{-/-}$ eye sections. Lenses are oriented with the anterior surface facing the upper right corner. Scale bar, 80 μ m. Abbreviations: e, anterior epithelial layer; f, lens fibre cells. The morphology of the $Rb^{+/+}$ lens was indistinguishable from the wild-type (data not shown).

METHODS. Embryos were fixed, sectioned parallel to the optic nerve, and stained with haematoxylin and eosin as described (S.D.M. et al., manuscript submitted). Embryos were genotyped from placental DNAs by PCR as described for the RB^{3kr} allele³ and the $p53^{\Delta}$ null allele¹⁹.

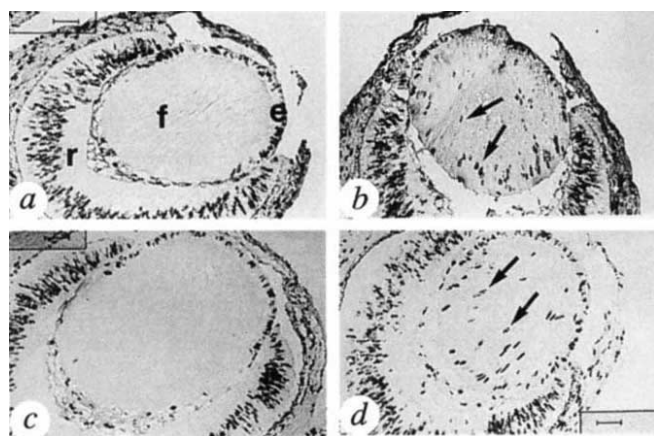


FIG. 2 Lens cell proliferation in E13.5 wild-type and mutant embryos. a, Wild-type; b, $Rb^{-/-}$; c, $p53^{-/-}$; and d, $Rb^{-/-}; p53^{-/-}$ eye sections. Orientation and scale bar as in Fig. 1. Arrows indicate representative peroxidase-stained, BrdU⁺ nuclei. Abbreviations as in Fig. 1, and r, retina. Nuclear-associated staining was not observed in the lens fibre cell compartment of $Rb^{+/+}$ embryos (data not shown).

METHODS. Embryos were exposed to BrdU *in utero*, fixed in formalin, embedded in paraffin, sectioned parallel to the optic nerve, and the sections assayed by indirect immunoperoxidase methods using an anti-BrdU/DNA antibody (S.D.M. et al., manuscript submitted). In control experiments, nuclear-associated staining was not detected when the anti-BrdU antibody was not included in the assay (data not shown).

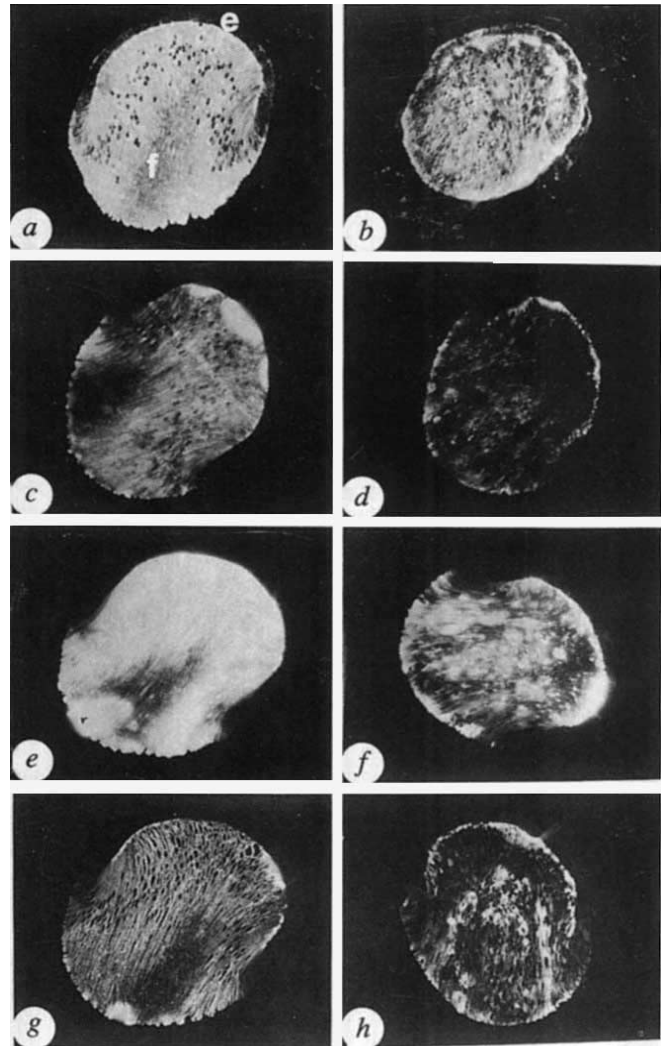


FIG. 3 Cellular differentiation in lenses derived from E14.5 wild-type and mutant embryos. Wild-type (a, c, e, g) and Rb -deficient (b, d, f, h) lenses were assayed for expression of α - (a, b), β - (c, d) and γ -crystallin (e, f), and MIP26 (g, h). Orientation and abbreviations as in Fig. 1. Magnification, $\times 37$. In the normal lens, α -crystallin is first observed at the lens vesicle stage and continues to be expressed in both epithelial and fibre cells⁹, whereas expression of β - and γ -crystallins and MIP26 begins at E12.5 and is restricted to differentiated lens fibre cells^{8,10}. Expression of the crystallins and MIP26 in $Rb^{+/+}$ lenses are indistinguishable from that in wild-type embryos (data not shown).

METHODS. Embryos were fixed, sectioned parallel to the optic nerve, and coronal sections assayed by indirect immunofluorescence with polyclonal antisera against α -, β - and γ -crystallin, and MIP26 (S.D.M. et al., manuscript submitted). Sections stained with the same antibody were photographed with equivalent exposure times. The specificity of these antibodies was confirmed by the absence of staining when the respective preimmune sera were used (data not shown).

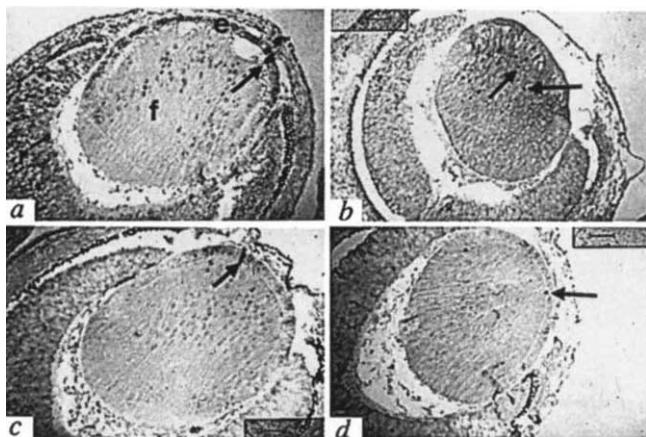


FIG. 4 Apoptotic cell death in lenses from E13.5 wild-type and mutant embryos. a, Wild-type; b, $Rb^{-/-}$; c, $p53^{-/-}$; and d, $Rb^{-/-}; p53^{-/-}$ eyes. Orientation and scale bar as in Fig. 1. Arrows indicate representative peroxidase-stained, TUNEL⁺ nuclei. The dark rim posterior to the retina is the pigmented layer, and is not due to the extension of free 3'-OH groups by the TUNEL assay. Abbreviations as in Fig. 2. b and d are representative lens sections of three $Rb^{-/-}$ and $p53^{-/-}$ embryos, respectively. To estimate the number of apoptotic cells per lens section in each embryo, peroxidase-stained nuclei were counted on five sections from the centremost region of the lens and were averaged. $Rb^{-/-}$ embryos averaged 12.4, 15.75 and 25.6 apoptotic nuclei per lens section each, whereas $Rb^{-/-}; p53^{-/-}$ embryos averaged 0.6, 1.0 and 1.6. To quantify the reduction in apoptosis in $Rb^{-/-}; p53^{-/-}$ lenses relative to $Rb^{-/-}$ lenses, the results for all three $Rb^{-/-}$ lenses and for all three $Rb^{-/-}; p53^{-/-}$ lenses were then averaged and compared. For $Rb^{-/-}$ and $Rb^{-/-}; p53^{-/-}$ lenses, the mean is 17.92 ± 5.6 and 1.07 ± 0.41 apoptotic nuclei per lens section, respectively, giving a reduction of 94%. Nuclear staining was not observed in the lens fibre cells compartment of $Rb^{+/+}$ embryos (data not shown).

METHODS. Embryos were fixed and coronal sections assayed for the presence of DNA free 3'-OH groups using a modified version of the TUNEL assay (S.D.M. et al., submitted). In control experiments, nuclear staining was absent when TdT and biotinylated deoxyuridine were excluded from the assay (data not shown).

or -independent pathways, even in the same cell type¹⁶⁻¹⁸. To determine whether apoptosis resulting from the lack of Rb in the lens was dependent on $p53$, we produced embryos deficient for both of these genes ($Rb^{-/-}; p53^{-/-}$; ref. 19) and examined lens development therein at the most advanced stage possible (E13.5-E14.0). The percentage of TUNEL-stained nuclei in $Rb^{-/-}; p53^{-/-}$ lenses was significantly diminished (by 94%) compared to that observed in age-matched $Rb^{-/-}$ embryos (Fig. 4b, d), indicating that apoptosis in Rb -deficient lens fibre cells was $p53$ -dependent. The reduction in apoptosis in double-mutant embryos did not correlate with any detectable changes in the rate of lens fibre cell proliferation (Fig. 2d), or in the degree of impaired morphological (Fig. 1d) and molecular differentiation (data not shown), nor were there any abnormalities in lens development in embryos deficient for $p53$ only (Figs 1c, 2c and 4c).

Our results establish a role for Rb in the development of the ocular lens. In particular, lens fibre cells lacking Rb do not withdraw from the cell cycle before the initiation of end-stage differentiation. These abnormally proliferating cells fail to express normal levels of late-stage differentiation markers and ultimately die by apoptosis. From the analysis of the most advanced lenses from embryos doubly mutant for Rb and $p53$, we can conclude that the death of the Rb -deficient lens fibre cells is $p53$ -dependent. This result is reminiscent of the requirement for $p53$ for apoptosis in cells expressing the adenovirus E1A oncoprotein^{14,15}, which sequesters member of the pRb family as well as other cellular proteins^{20,21}.

We have also observed abnormal development and extensive apoptosis in lenses of chimaeric mice composed of both Rb -deficient and wild-type cells, indicating that this process is lens-cell autonomous²². Moreover, we have shown that forced expression of a *c-myc* transgene also leads to deregulated lens fibre cell growth, but, in contrast to the effects noted here, these cells express normal levels of late-stage differentiation markers and do not undergo apoptosis (S.D.M. and R.A.D., manuscript submitted). Therefore, the pathological consequences of Rb deficiency may not be accounted for simply by ectopic mitotic activity.

The consequences of null Rb and $p53$ mutations on lens cell growth and death presented here have important implications for understanding the mechanism through which inactivation of their human counterparts may collaborate in human carcinogenesis. Many tumour types exhibit mutations in both Rb and $p53$ (ref. 19 and references therein) and mice that are $Rb^{+/-}; p53^{-/-}$ develop a wider range of tumours at earlier ages than mice that are either $Rb^{+/-}$ or $p53^{-/-}$ alone¹⁹, indicating that null mutations in both Rb and $p53$ can cooperate to bring about malignant transformation, and, that in some cell types, inactivation of both genes may be required. Moreover, the ability of several viruses to transform cells in culture²³⁻²⁵ and cause tumours in the mouse lens^{26,27} requires viral oncoproteins that bind to and inactivate both pRb and p53. Due to the attendant lethality of the $Rb^{-/-}; p53^{-/-}$ embryos, we were unable to observe tumorigenic consequences of these two mutations in the developing mouse lens.

Our results show that although loss of Rb is sufficient to deregulate lens cell proliferation, this activates a $p53$ -dependent programme that leads to the efficient elimination of these abnormally growing cells. This developmental study provides strong corroborating *in vivo* evidence that $p53$ may function as a tumour suppressor by promoting the death of pre-malignant cells which could eventually compromise the organism.

Note added in proof: Following submission of this paper, Pan and Griep²⁸ reported that expression of the pRb-binding protein HPV E7 in the postnatal mouse lens led to effects similar to those described here in the developing Rb -deficient lens. Moreover, co-expression of HPV E6 led to a reduction in apoptosis and to lens tumour formation. □

Received 31 May; accepted 19 July 1994.

- Hollingsworth, R. E., Hensley, C. E. & Lee, W. H. *Curr. Opin. Genet. Dev.* **3**, 55-62 (1993).
- Lee, E. Y.-H. P. et al. *Nature* **359**, 288-294 (1992).
- Jacks, T. et al. *Nature* **359**, 295-300 (1992).
- Clarke, A. R. et al. *Nature* **359**, 328-330 (1992).
- Pei, Y. F. & Rhodin, J. A. G. *Anat. Rec.* **169**, 105-126 (1970).
- Silver, J. & Hughes, A. F. W. *J. Morphol.* **140**, 159-170 (1973).
- Ishazaki, Y., Voyvodic, J. T., Burne, J. F. & Raff, M. C. *J. Cell Biol.* **121**, 899-908 (1993).
- McAvoy, J. W. *J. Embryol. exp. Morph.* **44**, 149-165 (1978).
- Shiels, A., Griffin, C. S. & Muggleton-Harris, A. L. *Biochim. biophys. Acta* **1097**, 318-324 (1991).
- Yancey, S. B., Koh, K., Chung, J. & Revel, J. P. *J. Cell Biol.* **106**, 705-714 (1988).
- Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. *J. Cell Biol.* **119**, 493-501 (1992).
- Yonish-Rouach, E. et al. *Nature* **352**, 345-347 (1991).
- Shaw, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4495-4499 (1992).
- Debbas, M. & White, E. *Genes Dev.* **7**, 546-554 (1993).
- Lowe, S. W., Jacks, T., Housman, D. E. & Ruley, H. E. *Proc. natn. Acad. Sci. U.S.A.* **91**, 2026-2030 (1994).
- Lowe, S. W. et al. *Nature* **362**, 847-849 (1993).
- Clarke, A. R. et al. *Nature* **362**, 849-852 (1993).
- Berges, R. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8910-8914 (1993).
- Williams, B. O. et al. *Nature Genet.* **7**, 480-484 (1994).
- Whyte, P. et al. *Nature* **334**, 124-129 (1988).
- Pines, J. & Hunter, T. *Nature* **346**, 760-763 (1990).
- Williams, B. O. et al. *EMBO J.* (in press).
- Hawley-Nelson, P. et al. *EMBO J.* **8**, 3905-3910 (1989).
- Watanabe, S., Kanda, T. & Yoshiike, K. *J. Virol.* **63**, 965-969 (1989).
- Munger, K. et al. *J. Virol.* **63**, 4417-4421 (1989).
- Mahon, K. et al. *Science* **235**, 1622-1628 (1987).
- Griep, A. et al. *J. Virol.* **67**, 1373-1384 (1993).
- Pan, H. & Griep, A. E. *Genes Dev.* **8**, 1285-1299 (1994).

ACKNOWLEDGEMENTS. We thank S. Zigler and J. Horwitz for antibodies against the crystallins and MIP26 protein; K. Macleod, S. Lowe and L. Remington for advice and assistance; and N. Schreiber-Agus for critically reading the manuscript. S.D.M. and B.O.W. were supported by NIH predoctoral training grants. This work was supported by grants from the National Eye Institute and National Institute for Child Health and Human Development and an NCI-Cancer core grant (R.A.D.) as well as the Searle Scholars Program/The Chicago Community Trust and NCI (T.I.).

Mutations of two *PMS* homologues in hereditary nonpolyposis colon cancer

Nicholas C. Nicolaides*, Nickolas Papadopoulos*, Bo Liu*, Ying-Fei Wei†, Kenneth C. Carter†, Steven M. Ruben‡, Craig A. Rosen‡, William A. Haseltine‡, Robert D. Fleischmann‡, Claire M. Fraser‡, Mark D. Adams‡, J. Craig Venter‡, Malcolm G. Dunlop§, Stanley R. Hamilton||, Gloria M. Petersen¶, Albert de la Chapelle#, Bert Vogelstein* & Kenneth W. Kinzler*††

* The Johns Hopkins Oncology Center, Baltimore, Maryland 21231, USA

† Human Genome Sciences, Inc., 9620 Medical Center Drive, Suite 300, Rockville, Maryland 20850-3338, USA

‡ The Institute for Genomic Research, 932 Clopper Road, Gaithersburg, Maryland 20878, USA

§ MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK

|| The Johns Hopkins University School of Medicine, Departments of Pathology and Oncology, Baltimore, Maryland 21205, USA

¶ Department of Epidemiology,

The Johns Hopkins University School of Public Health and Hygiene, Baltimore, Maryland 21205, USA

Department of Medical Genetics, University of Helsinki, 00290 Helsinki, Finland

HEREDITARY nonpolyposis colorectal cancer (HNPCC) is one of man's commonest hereditary diseases¹. Several studies have implicated a defect in DNA mismatch repair in the pathogenesis of this disease²⁻⁸. In particular, *hMSH2* and *hMLH1* homologues of the bacterial DNA mismatch repair genes *mutS* and *mutL*, respectively, were shown to be mutated in a subset of HNPCC cases⁹⁻¹⁶. Here we report the nucleotide sequence, chromosome localization and mutational analysis of *hPMS1* and *hPMS2*, two additional homologues of the prokaryotic *mutL* gene. Both *hPMS1* and *hPMS2* were found to be mutated in the germline of HNPCC patients. This doubles the number of genes implicated in HNPCC and may help explain the relatively high incidence of this disease.

We previously screened a database of human genes identified by the expressed sequence tag (EST) method¹⁷ and identified three ESTs with homology to bacterial and yeast *mutL* genes¹⁶. Two ESTs had greatest homology to the yeast *mutL*-homologue *yPMS1*, and therefore were designated *hPMS1* and *hPMS2*. Sequence analysis of the *hPMS1* complementary DNA clone identified an open reading frame (ORF) of 2,795 basepairs (bp) flanked on both sides by in-frame termination codons (Fig. 1a). If initiation occurred at the first methionine of the ORF, a 932 residue protein with 27% identity to *yPMS1* would be produced. Sequence analysis of the *hPMS2* clone identified a 2,586 bp ORF (Fig. 1b) without an upstream stop codon. Attempts to extend the 5' end sequence using the rapid amplification of cDNA ends (RACE) technique¹⁸, combined with genomic sequencing revealed a complex pattern suggesting alternative initiation sites (data not shown). Sequence analysis of the major 5' products failed to reveal an upstream in-frame stop codon. However, the methionine at nucleotide 25 fit the context of a Kozak consensus sequence for translation initiation at 5 of 7 positions¹⁹. If initiation occurred at this methionine, an 862 residue protein with 32% identity to *yPMS1* would be produced. Comparison of the yeast and human PMS proteins showed that the greatest homo-

logy was within the amino-terminal segment, from codons 40 to 390, which was 30% identical between *yPMS1* and *hPMS1*, and 42% between *yPMS1* and *hPMS2* (Fig. 1c). This region contained several domains that are highly conserved among the *mutL*-related proteins, most notably the GFRGEAL domain which is perfectly conserved in *E. coli* and human *mutL* homologues²⁰. A second region of significant homology was in the carboxyl terminus. This region was 22% identical between *yPMS1* and *hPMS1* and 47% identical between *yPMS1* and *hPMS2*, while very little homology was observed in this region between these proteins and the yeast *mutL*-homologue *yMLH1* (not shown).

hPMS1 and *hPMS2* were previously localized to chromosomes 2 and 7, respectively¹⁶. To determine the precise location of these genes, genomic clones were used for fluorescent in situ hybridization (FISH) to human metaphase chromosome spreads. *hPMS1* was localized to chromosome 2q31-33 using a genomic P1 clone (1670) which contained its 5' end (Fig. 2a, b, c). Likewise, the *hPMS2* gene was localized to chromosome 7p22 using a genomic P1 clone (2053) which contained the 3' region of the *hPMS2* gene (Fig. 2d, e). Liskay and colleagues have also cloned a human *mutL* homologue which maps to chromosome 7 (R. M. Liskay, personal communication)¹⁵. Analysis with a variety of genomic clones indicates that *hPMS2* is a member of a subfamily which includes at least two related genes located on chromosome 7q.

To evaluate the role of *hPMS1* and *hPMS2* in HNPCC, we examined 40 patients with a family history of HNPCC. Eighteen were found to have mutations of *hMSH2* or *hMLH1* and therefore were not fully evaluated for *hPMS1* or *hPMS2* mutations^{13,16,21}. To identify potential mutations in the other 22 samples, we employed an *in vitro* synthesized protein (IVSP) assay which detects deletions, insertions, frameshifts and nonsense mutations that result in an altered size of the protein product^{22,23}. Three overlapping segments of *hPMS1* were amplified using reverse transcriptase polymerase chain reaction (RT-PCR) and the products were transcribed and translated *in vitro*. All samples generated a protein of the expected size, although a reproducible truncated product was detected in a segment containing codons 1 to 500 from patient C.W. The expected relative molecular mass of the wild-type protein from this segment was 55,000 (55 K) while the truncated protein product had an approximate molecular weight of 25 K. Sequence analysis of RT-PCR products from this region revealed an in-frame deletion that removed codons 195 to 233. Because this small deletion could not result in the 25 K truncated protein, we attempted to map the truncation by amplifying cDNA from C.W. using the same 5' primer and nested 3' primers. IVSP analysis revealed that the truncated polypeptide was present in samples translated from templates containing codons 1 to 369 and 1 to 290 but was absent in a template containing codons 1 to 214 (Fig. 3a). This placed the potential stop codon mutation between codons 214 and 290, in a region overlapping with the in-frame deletion. As it has been previously noted that exons containing nonsense codons can be selectively skipped²⁴, we determined the intron-exon boundaries in this region. Sequence analysis of an *hPMS1* genomic clone identified an exon which corresponded exactly to the boundaries of the deletion (codons 195-233). We then amplified the skipped exon from the genomic DNA of patient C.W. While no alterations of the splicing consensus sites were identified, a single C to T transition at codon 233, which converted a glutamine to stop codon, was identified (Fig. 3b, lane 2). This nonsense codon appeared to be responsible for the exon-skipping.

For mutational analysis of *hPMS2*, cDNA was amplified either as a full-length product or as two overlapping segments and used in the IVSP assay. In one case (G.C.), analysis of the full-length product identified an additional truncated protein of about 50 K (Fig. 4a, lane GC). Moreover, analysis of *hPMS2* in a tumour xenograft from this patient revealed the 50 K truncated

†† To whom correspondence should be addressed

a

1 GG CAC GAG TGG CTG CTT GCG GCT AGT GGA TGG TAA TTG CCT GCC TCG CGE TAG CAG CAA GCT GCT CTG TTA AAA GCG AAA ATG Met

2 Lys Gln Leu Pro Ala Ala Thr Val Arg Leu Leu Ser Ser Ser Gln Ile Ile Thr Ser Val Val Ser Val Val Lys Glu Leu Ile

84 AAA CAA TTG CCT GCG GCA ACA GTT CGA CTC CTT TCA AGT TCT CAG ATC ATC ACT TCG GTG GTC AGT GTT GTA AAA GAG CTT ATT

30 Glu Asn Ser Ser Leu Asp Ala Gly Ala Thr Ser Val Asp Val Lys Leu Glu Asn Tyr Gly Phe Asp Lys Ile Glu Val Arg Asp Asn

168 GAA AAC TCC TTG GAT GCT GGT GCC ACA AGC GTA GAT GTT AAA CTG GAG AAC TAT GGA TTT GAT AAA ATT GAG GTG CGA GAT AAC

58 Gly Glu Gly Ile Lys Ala Val Asp Ala Pro Val Met Ala Met Lys Tyr Tyr Thr Ser Lys Ile Asn Ser His Glu Asp Leu Glu

252 GGG GAG GGT ATC AAG GCT GTT GAT GCA CCT GTA ATG GCA ATG AAG TAC TAC ACC TCA AAA ATA AAT AGT CAT GAA GAT CTT GAA

86 Asn Leu Thr Thr Tyr Gly Phe Arg Gly Glu Ala Leu Gly Ser Ile Cys Cys Ile Ala Glu Val Leu Ile Thr Thr Arg Thr Ala

336 AAT TTG ACA ACT TAC GGT TTT CGT GGA GAA GCC TTG GGG TCA ATT TGT TGT ATA GCT GAG GTT TTA ATT ACA ACA AGA ACG GCT

114 Ala Asp Asn Phe Ser Thr Gln Tyr Val Leu Asp Gly Ser Gly His Ile Leu Ser Gln Lys Pro Ser His Leu Gly Gln Gly Thr

420 GCT GAT AAT TTT AGC ACC CAG TAT GTT TTA GAT GGC AGT GGC CAC ATA CTT TCT CAG AAA CCT TCA CAT CTT GGT CAA GGT ACA

142 Thr Val Thr Ala Leu Arg Leu Phe Lys Asn Leu Pro Val Arg Lys Gln Phe Tyr Ser Thr Ala Lys Lys Cys Lys Asp Glu Ile

504 ACT GTA ACT GCT TTA AGA TTA TTT AAG AAT CTA CCT GTA AGA AAG CAG TTT TAC TCA ACT GCA AAA AAA TGT AAA GAT ACA ATA

170 Lys Lys Ile Gln Asp Leu Leu Met Ser Phe Gly Ile Leu Lys Pro Asp Leu Arg Ile Val Phe Val His Asn Lys Ala Val Ile

588 AAA AAG ATC CAA GAT CTC CTC ATG AGC TTT GGT ATC CTT AAA CCT GAC TTA AGG ATT GTC TTT GAT CAT AAC AAG GCA GTT ATT

198 Trp Gln Lys Ser Arg Val Ser Asp His Lys Met Ala Leu Met Ser Val Leu Gly Thr Ala Val Met Asn Asn Met Glu Ser Phe

672 TGG CAG AAA AGC AGA GTA TCA GAT CAC AAG ATG GCT CTC ATG TCA GTT CTG GGG ACT GCT ATA AAT GAT AAA GAT AAA TCC TTT

226 Gln Tyr His Ser Glu Glu Ser Gln Ile Tyr Leu Ser Gly Phe Leu Pro Lys Cys Asp Ala Asp His Ser Phe Thr Ser Leu Ser

756 CAG TAC CAC TCT GAA GAA TCT CAG ATT TAT CTC AGT GGA TTT CTT CCA AAG TGT GAT GCA GAC CAC TCT TTC ACT AGT CTT TCA

254 Thr Pro Glu Arg Ser Phe Ile Phe Ile Asn Ser Arg Pro Val His Gln Lys Asp Ile Leu Lys Leu Ile Arg His His Tyr Asn

840 ACA CCA GAA AGA AGT TTC ATC TTC ATA AAC AGT CGA CCA GTA CAT CAA AAA GAT ATC TTA AAG TTA ATC CGA CAT CAT TAC AAT

282 Leu Lys Cys Leu Lys Glu Ser Thr Arg Leu Tyr Pro Val Phe Phe Leu Lys Ile Asp Val Pro Thr Ala Asp Val Asp Val Asn

924 CTG AAA TGC CTA AAG GAA TCT ACT CGT TTG TAT CCT GTT TTC TTT CTG AAA ATC GAT GTT CCT ACA GCT GAT GTT GAT GTA AAT

310 Leu Thr Pro Asp Lys Ser Gln Val Leu Leu Gln Asn Lys Glu Ser Val Leu Ile Ala Leu Glu Asn Leu Met Thr Thr Cys Tyr

1008 TTA ACA CCA GAT AAA AGC CAA GTA TTA TTA CAA AAT AAG GAA TCT GTT TTA ATT GCT CTT GAA AAT CTG ATG ACG ACT TGT TAT

338 Gly Pro Leu Pro Ser Thr Asn Ser Tyr Glu Asn Asn Lys Thr Asp Val Ser Ala Ala Asp Ile Val Leu Ser Lys Thr Ala Glu

1092 GGA CCA TTA CCT AGT ACA AAT TCT TAT GAA AAT AAT AAA ACA GAT GTT TCC GCA GCT GAC ATC GTT CTT AGT AAA ACA GCA GAA

366 Thr Asp Val Leu Phe Asn Lys Val Glu Ser Ser Gly Lys Asn Tyr Ser Asn Val Asp Thr Ser Val Ile Pro Phe Gln Asn Asp

1176 ACA GAT GTG CTT TTT AAT AAA GTG GAA TCA TCT GGA AAG AAT TAT TCA AAT GTT GAT ACT TCA GTC ATT CCA TTC CAA AAT GAT

394 Met His Asn Asp Glu Ser Gly Lys Asn Thr Asp Asp Cys Leu Asn His Gln Ile Ser Ile Gly Asp Phe Gly Tyr Gly His Cys

1260 ATG CAT AAT GAT GAA TCT GGA AAA AAC ACT GAT GAT TGT TTA AAT CAC CAG ATA AGT AAT GGT GAC TTT GGT TAT GGT CAT TGT

422 Ser Ser Glu Ile Ser Asn Ile Asp Lys Asn Thr Lys Asn Ala Phe Gln Asp Ile Ser Met Ser Asn Val Ser Trp Glu Asn Ser

1344 AGT AGT GAA ATT TCT AAC ATT GAT AAA AAC ACT AAG AAT GCA TTT CAG GAC ATT TCA ATG AGT AAT GTA TCA TGG GAG AAC TCT

450 Gln Thr Glu Tyr Ser Lys Thr Cys Phe Ile Ser Ser Val Lys His Thr Gln Ser Glu Asn Gly Asn Lys Asp His Ile Asp Glu

1428 CAG ACG GAA TAT AGT AAA ACT TGT TTT ATA AGT TCC GTT AAG CAC ACC CAG TCA GAA AAT GGC AAT AAA GAT CAT ATA GAT GAT

478 Ser Gly Glu Asn Glu Glu Glu Ala Gly Leu Glu Asn Ser Ser Glu Ile Ser Ala Asp Glu Trp Ser Arg Gly Asn Ile Leu Lys

1512 AGT GGG GAA AAT GAG GAA GAA GCA GGT CTT GAA AAC TCT TCG GAA ATT TCT GCA GAT GAG TGG AGC AGG GGA AAT ATA CTT AAA

506 Asn Ser Val Gly Glu Asn Ile Glu Pro Val Lys Ile Leu Val Pro Glu Lys Ser Leu Pro Cys Lys Val Ser Asn Asn Asn Tyr

1596 AAT TCA GTG GGA GAG AAT ATT GAA CCT GTG AAA ATT TTA GTG CCT GAA AAA AGT TTA CCA TGT AAA GTA AGT AAT AAT AAT TAT

534 Pro Ile Pro Glu Gln Met Asn Leu Asn Glu Asp Ser Cys Asn Lys Lys Ser Asn Val Ile Asp Asn Lys Ser Gly Lys Val Thr

1680 CCA ATC CCT GAA CAA ATG AAT CTT AAT GAA GAT TCA TGT AAC AAA AAA TCA AAT GTA ATA GAT AAT AAA TCT GGA AAA GTT ACA

562 Ala Tyr Asp Leu Leu Ser Asn Arg Val Ile Lys Lys Pro Met Ser Ala Ser Ala Leu Phe Val Gln Asp His Arg Pro Gln Phe

1764 GCT TAT GAT TTA CTT AGC AAT CGA GTA ATC AAG AAA CCC ATG TCA GCA AGT GCT CTT TTT GTT CAA GAT CAT CGT CCT CAG TTT

590 Leu Ile Glu Asn Pro Lys Thr Ser Leu Glu Asp Ala Thr Ser Leu Glu Lys Thr Leu Thr Ser Leu Thr Ser Leu Thr Ser Leu Thr

1848 CTC ATA GAA AAT CCT AAG ACT AGT TTA GAG GAT GCA ACA CTA CAA ATT GAA GAA CTG TGG AAG ACA TTG AGT GAA GAG GAA AAA

618 Leu Lys Tyr Glu Glu Lys Ala Thr Lys Asp Leu Glu Arg Tyr Asn Ser Gln Met Lys Arg Ala Ile Glu Gln Glu Ser Gln Met

1932 CTG AAA TAT GAA GAG AAG GCT ACT AAA GAC TTG GAA CGA TAC AAT AGT CAA ATG AAG AGA GCC ATT GAA CAG GAG TCA CAA ATG

646 Ser Leu Lys Asp Gly Arg Lys Lys Ile Lys Pro Thr Ser Ala Trp Asn Leu Ala Gln Lys His Lys Leu Lys Thr Ser Leu Thr

2016 TCA CTA AAA GAT GGC AGA AAA AAG ATA AAA CCC ACC AGC GCA TGG AAT TTG GCC CAG AAG CAC AAG TTA AAA ACC TCA TTA TCT

674 Asn Gln Pro Lys Leu Asp Glu Leu Leu Gln Ser Gln Ile Glu Lys Arg Arg Ser Gln Asn Ile Lys Met Val Gln Ile Pro Phe

2100 AAT CAA CCA AAA CTT GAT GAA CTC CTT CAG TCC CAA ATT GAA AAA AGA AGG AGT CAA AAT ATT AAA ATG GTA CAG ATC CCC TTT

702 Ser Met Lys Asn Leu Lys Ile Asn Phe Lys Lys Gln Asn Lys Val Asp Leu Glu Glu Lys Asp Glu Pro Cys Leu Ile His Asn

2184 TCT ATG AAA AAC TTA AAA ATA AAT TTT AAG AAA CAA AAC AAA GTT GAC TTA GAA GAG AAG GAT GAA CCT TGC TTG ATC CAC AAT

730 Leu Arg Phe Pro Asp Ala Trp Leu Met Thr Ser Lys Thr Glu Val Met Leu Leu Asn Pro Tyr Arg Val Glu Glu Ala Leu Leu

2268 CTC AGG TTT CCT GAT GCA TGG CTA ATG ACA TCC AAA ACA GAG GTA ATG TTA TTA AAT CCA TAT AGA GTA GAA GAA GCC CTG CTA

758 Phe Lys Arg Leu Thr Glu Asn His Lys Leu Pro Cys Glu Glu Pro Leu Glu Lys Pro Ile Met Leu Thr Glu Ser Leu Phe Asn Gly

2352 TTT AAA AGA CTT CTT GAG AAT CAT AAA CTT CCT GCA GAG CCA CTG GAA AAG CCA ATT ATG TTA ACA GAG AGT CTT TTT AAT GGA

786 Ser His Tyr Leu Asp Val Leu Tyr Lys Met Thr Ala Asp Asp Gln Arg Tyr Ser Gly Ser Thr Tyr Leu Ser Asp Pro Arg Leu

2436 TCT CAT TAT TTA GAC GTT TTA TAT AAA ATG ACA GCA GAT GAC CAA AGA TAC AGT GGA TCA ACT TAC CTG TCT GAT CCT CGT CTT

814 Thr Ala Asn Gly Phe Lys Ile Lys Leu Ile Pro Gly Val Ser Ile Thr Glu Asn Tyr Leu Glu Ile Glu Gly Met Ala Asn Cys

2520 ACA CCG AAT GGT TTC AAG ATA AAA TTG ATA CCA GGA GTT TCA ATT ACT GAA AAT TAC TTG GAA ATA GAA GGA ATG GCT AAT TGT

842 Leu Pro Phe Tyr Gly Val Ala Asp Leu Lys Glu Ile Leu Asn Ala Ile Leu Asn Arg Asn Ala Lys Glu Val Tyr Glu Cys Arg

2604 CTC CCA TTC TAT GGA GTA GCA GAT TTA AAA GAA ATT CTT AAT GCT ATA TTA AAC AGA AAT GCA AAG GAA GTT TAT GAA TGT AGA

870 Pro Arg Lys Val Ile Ser Tyr Leu Glu Gly Glu Ala Val Arg Leu Ser Arg Gln Leu Pro Met Tyr Leu Ser Lys Glu Asp Ile

2688 CCT CGC AAA GTG ATA AGT TAT TTA GAG GGA GAA GCA GTG CGT CTA TCC AGA CAA TTA CCC ATG TAC TTA TCA AAA GAG GAC ATC

898 Gln Asp Ile Ile Tyr Arg Met Lys His Gln Phe Gly Asn Glu Ile Lys Glu Cys Val His Gly Arg Pro Phe Phe His His Leu

2772 CAA GAC ATT ATC TAC AGA ATG AAG CAC CAG TTT GGA AAT GAA ATT AAA GAG TGT GTT CAT GGT CGC CCA TTT TTT CAT CAT TTA

926 Thr Tyr Leu Pro Glu Thr Thr

2856 ACC TAT CTT CCA GAA ACT ACA TGA TTA AAT ATG TTT AAG AAG ATT AGT TAC CAT TGA AAT TGG TTC TGT CAT AAA ACA GCA TGA

2940 GTC TGG TTT TAA ATT ATC TTT GTA TTA TGT GTC ACA TGG TTA TTT TTT AAA TGA GGA TTC ACT GAC TTG TTT TTA TAT TGA AAA

3024 AAG TTC CAC GTA TTG TAG AAA ACG TAA ATA AAC TAA TAA C

FIG. 1 Structure of human *hPMS1* and *hPMS2*. The cDNA sequence and predicted amino acid sequence of the protein encoded by *hPMS1* (a) and *hPMS2* (b) are shown. c (shown on page 78), Alignment of the predicted amino-acid sequences of the *S. cerevisiae* PMS1 (*yPMS1*)²⁶, the human PMS1 (*hPMS1*) and the human PMS2 (*hPMS2*) proteins using MACAW (version 1.0) program. Amino acids in conserved blocks are capitalized and shaded based on the mean of their pair wise scores²⁷. The *hPMS1* and *hPMS2* nucleotide sequences have been deposited with Genbank (accession numbers U13695 and U13696, respectively).

b
1
1 CGA GGC GGA TCG GGT GTT GCA TCC ATG GAG CGA GCT GAG AGC TCG AGT ACA GAA CCT GCT AAG GCC ATC AAA CCT ATT GAT CGG
21 Lys Ser Val His Gln Ile Cys Ser Gly Gln Val Val Leu Ser Leu Ser Thr Ala Val Lys Glu Leu Val Glu Asn Ser Leu Asp
85 AAG TCA GTC CAT CAG ATT TGC TCT GGG CAG GTG GTA CTG AGT CTA AGC ACT GCG GTA AAG GAG TTA GTA GAA AAC AGT CTG GAT
49 Ala Gly Ala Thr Asn Ile Asp Leu Lys Leu Lys Asp Tyr Gly Val Asp Leu Ile Glu Val Ser Asp Asn Gly Cys Gly Val Glu
169 GCT GGT GCC ACT AAT ATT GAT CTA AAG CTT AAG GAC TAT GGA GTG GAT CTT ATT GAA GTT TCA GAC AAA ATG GTC CAG GTC GAA
77 Glu Glu Asn Phe Glu Gly Leu Thr Leu Lys His His Thr Ser Lys Ile Gln Glu Phe Ala Asp Leu Thr Gln Val Glu Thr Phe
253 GAA GAA AAC TTC GAA GGC TTA ACT CTG AAA CAT CAC ACA TCT AAG ATT CAA GAG TTT GCC GAC CTA ACT CAG GTT GAA ACT TTT
105 Gly Phe Arg Gly Glu Ala Leu Ser Ser Leu Cys Ala Leu Ser Asp Val Thr Ile Ser Thr Cys His Ala Ser Ala Lys Val Gly
337 GGC TTT CCG GGG GAA GCT CTG AGC TCA CTT TGT GCA CTG AGC GAT GTC ACC ATT TCT ACC TGC CAC GCA TCG GCG AAG GTT GGA
133 Thr Arg Leu Met Phe Asp His Asn Gly Lys Ile Ile Gln Lys Thr Pro Tyr Pro Arg Pro Arg Gly Thr Thr Val Ser Val Gln
421 ACT CGA CTG ATG TTT GAT CAC AAT GGG AAA ATT ATC CAG AAA ACC CCC TAC CCC CGC CCC AGA GGG ACC ACA GTC AGC GTG CAG
161 Gln Leu Phe Ser Thr Leu Pro Val Arg Ser Arg His Lys Glu Phe Gln Arg Asn Ile Lys Lys Glu Tyr Ala Lys Met Val Gln Val Leu
505 CAG TTA TTT TCC ACA CTA CTT GTG CGC CAT AAG GAA TTT CAA AGG AAT ATT AAG AAG GAG TAT GCC AAA ATG GTC CAG GTC TTA
189 His Ala Tyr Cys Ile Ile Ser Ala Gly Ile Arg Val Ser Cys Thr Asn Gln Leu Gly Gln Gly Lys Arg Gln Pro Val Val Cys
589 CAT GCA TAC TGT ATC ATT TCA GCA GGC ATC CGT GTA AGT TGC ACC AAT CAG CTT GGA CAA GGA AAA CGA CAG CCT GTG GTA TGC
217 Thr Gly Gly Ser Pro Ser Ile Lys Glu Asn Ile Gly Ser Val Phe Gly Gln Lys Gln Leu Gln Ser Leu Ile Pro Phe Val Gln
673 ACA GGT GGA AGC CCC AGC ATA AAG GAA AAT ATC GGC TCT GTG TTT GGG CAG AAG CAG TTG CAA AGC CTC ATT CCT TTT GTT CAG
245 Leu Pro Pro Ser Asp Ser Val Cys Glu Glu Tyr Gly Leu Ser Cys Ser Asp Ala Leu His Asn Leu Phe Tyr Ile Ser Gly Phe
757 CTG CCC CTT AGG GAC TCC GTG TGT GAA GAG TAC GGT TTG AGC TGT TCG GAT GCT CTG CAT AAT CTT TAT TAC ATC TCA GGT TTT
273 Ile Ser Gln Cys Thr His Gly Val Gly Arg Ser Ser Thr Asp Arg Gln Phe Phe Phe Ile Asn Arg Arg Pro Cys Asp Pro Ala
841 ATT TCA CAA TGC ACG CAT GGA GTT GGA AGG AGT TCA ACA GAC AGA CAG TTT TTC TTT ATC AAC CGG CGG CCT TGT GAC CCA GCA
301 Lys Val Cys Arg Leu Val Asn Glu Val Tyr His Met Tyr Asn Arg His Gln Tyr Pro Phe Val Val Leu Asn Ile Ser Val Asp
925 AAG GTC TGC AGA CTC GTG AAT GAG GTC TAC CAC ATG TAT AAT CGA CAC CAG TAT CCA TTT GTT GTT CTT AAC ATT TCT GTT GAT
329 Ser Glu Cys Val Asp Ile Asn Val Thr Pro Asp Lys Arg Gln Ile Leu Leu Gln Glu Glu Lys Leu Leu Leu Ala Val Leu Lys
1009 TCA GAA TGC GTT GAT ATC AAT GTT ACT CCA GAT AAA AGG CAA ATT TTG CTA CAA GAG GAA AAG CTT TTG TTG GCA GTT TTA AAG
357 Thr Ser Leu Ile Gly Met Phe Asp Ser Asp Val Asn Lys Leu Asn Val Ser Gln Gln Pro Leu Leu Asp Val Glu Gly Asn Leu
1093 ACC TCT TTG ATA GGA ATG TTT GAT AGT GAT GTC AAC AAG CTA AAT GTC AGT CAG CAG CCA CTG CTG GAT GTT GAA GGT AAC TTA
385 Ile Lys Met His Ala Ala Asp Leu Glu Lys Pro Met Val Glu Lys Gln Asp Gln Ser Pro Ser Leu Arg Thr Gly Glu Glu Lys
1177 ATA AAA ATG CAT GCA GCG GAT TTG GAA AAG CCC ATG GTA GAA AAG CAG GAT CAA TCC CCT TCA TTA AGG ACT GGA GAA GAA AAA
413 Lys Asp Val Ser Ile Ser Arg Leu Arg Glu Ala Phe Ser Leu Arg His Thr Thr Glu Asn Lys Pro His Ser Pro Lys Thr Pro
1261 AAA GAC GTG TCC ATT TCC AGA CTG CGA GAG GCC TTT TCT CTT CGT CAC ACA ACA GAG AAC AAG CCT CAC AGC CCA AAG ACT CCA
441 Glu Pro Arg Arg Ser Pro Leu Gly Gln Lys Arg Gly Met Leu Ser Ser Ser Thr Ser Gly Ala Ile Ser Asp Lys Gly Val Leu
1345 GAA CCA AGA AGG AGC CCT CTA GGA CAG AAA AGG GGT ATG CTG TCT TCT AGC ACT TCA GGT GCC ATC TCT GAA GGC GTG CTG
469 Arg Pro Gln Lys Glu Ala Val Ser Ser Ser His Gly Pro Ser Asp Pro Thr Asp Arg Ala Glu Val Glu Lys Asp Ser Gly His
1429 AGA CCT CAG AAA GAG GCA GTG AGT TCC AGT CAC GGA CCC AGT GAC CCT ACG GAC AGA GCG GAG GTG GAG AAG GAC TCG GGG CAC
497 Gly Ser Thr Ser Val Asp Ser Glu Gly Phe Ser Ile Pro Asp Thr Gly Ser His Cys Ser Ser Glu Tyr Ala Ala Ser Ser Pro
1513 GGC AGC ACT TCC GTG GAT TCT GAG GGG TTC AGC ATC CCA GAC ACG GGC AGT CAC TGC AGC AGC GAG TAT GCG GCC AGC TCC CCA
525 Gly Asp Arg Gly Ser Gln Glu His Val Asp Ser Gln Glu Lys Ala Pro Glu Thr Asp Asp Ser Phe Ser Asp Val Asp Cys His
1597 GGG GAC AGG GGC TCG CAG GAA CAT GTG GAC TCT CAG GAG AAA GCG CCT GAA ACT GAC GAC TCT TTT TCA GAT GTG GAC TGC CAT
553 Ser Asn Gln Glu Asp Thr Gly Cys Lys Phe Arg Val Leu Pro Gln Pro Thr Asn Leu Ala Thr Pro Asn Thr Lys Arg Phe Lys
1681 TCA AAC CAG AAG GAT ACC GGA TGT AAA TTT CGA GTT TGT CCT CAG CCA ACT AAT CTC GCA ACC CCA AAC ACA AAG CGT TTT AAA
581 Lys Glu Glu Ile Leu Ser Ser Ser Asp Ile Cys Gln Lys Leu Val Asn Thr Gln Asp Met Ser Ala Ser Gln Val Asp Val Ala
1765 AAA GAA GAA ATT CTT TCC AGT TCT GAC ATT TGT CAA AAG TTA GTA AAT ACT CAG GAC ATG TCA GCC TCT CAG GTT GAT GTA GCT
609 Val Lys Ile Asn Lys Lys Val Val Pro Leu Asp Phe Ser Met Ser Ser Leu Ala Lys Arg Ile Lys Gln Leu His His Glu Ala
1849 GTG AAA ATT AAT AAG AAA GTT GTG CCC CTG GTC TTT TCT ATG AGT TCT TTA GCT AAA CGA ATA AAG CAG TTA CAT CAT GAA GCA
637 Gln Gln Ser Glu Gly Glu Gln Asn Tyr Arg Lys Phe Arg Ala Lys Ile Cys Pro Gly Glu Asn Gln Ala Ala Glu Asp Glu Leu
1933 CAG CAA AGT GAA GGG GAA CAG AAT TAC AGG AAG TTT AGG GCA AAG ATT TGT CCT GGA GAA AAT CAA GCA GCC GAA GAT GAA CTA
665 Arg Lys Glu Ile Ser Lys Thr Met Phe Ala Glu Met Glu Ile Ile Gly Gln Phe Asn Leu Gly Phe Ile Ile Thr Lys Leu Asn
2017 AGA AAA GAG ATA AGT AAA ACG ATG TTT GCA GAA ATG GAA ATC ATT GGT CAG TTT AAC CTG GGA TTT ATA ATA ACC AAA CTG AAT
693 Glu Asp Ile Phe Ile Val Asp Gln His Ala Thr Asp Glu Lys Tyr Asn Phe Glu Met Leu Gln Gln His Thr Val Leu Gln Gly
2101 GAG GAT ATC TTC ATA GTG GAC CAG CAT GCC ACG GAC GAG AAG TAT AAC TTC GAG ATG CTG CAG CAG CAC ACC GTG CTC CAG GGG
721 Gln Arg Leu Ile Ala Pro Gln Thr Leu Asn Leu Thr Ala Val Asn Glu Ala Val Leu Ile Glu Asn Leu Glu Ile Phe Arg Lys
2185 CAG AGG CTC ATA GCA CCT CAG ACT CTC AAC TTA ACT GCT GTT AAT GAA GCT GTT CTG ATA GAA AAT CTG GAA ATA TTT AGA AAG
749 Asn Gly Phe Asp Phe Val Ile Asp Glu Asn Ala Pro Val Thr Glu Arg Ala Lys Leu Ile Ser Leu Pro Thr Ser Lys Asn Trp
2269 AAT GGC TTT GAT TTT GTT ATC GAT GAA AAT GCT CCA GTC ACT GAA AGG GCT AAA CTG ATT TCC TTG CCA ACT AGT AAA AAC TGG
777 Thr Phe Gly Pro Gln Asp Val Asp Glu Leu Ile Phe Met Leu Ser Asp Ser Pro Gly Val Met Cys Arg Pro Ser Arg Val Lys
2353 ACC TTC GGA CCC CAG GAC GTC GAT GAA CTG ATC TTC ATG CTG AGC GAC AGC CCT GGG GTC ATG TGC CGG CCT TCC CGA GTC AAG
805 Gln Met Phe Ala Ser Arg Ala Cys Arg Lys Ser Val Met Ile Gly Thr Ala Leu Asn Thr Ser Glu Met Lys Lys Leu Ile Thr
2437 CAG ATG TTT GCC TCC AGA GCC TGC CGG AAG TCG GTG ATG ATT GGG ACT GCT CTT AAC ACA AGC GAG ATG AAG AAA CTG ATC ACC
833 His Met Gly Glu Met Asp His Pro Trp Asn Cys Pro His Gly Arg Pro Thr Met Arg His Ile Ala Asn Leu Gly Val Ile Ser
2521 CAC ATG GGG GAG ATG GAC CAC CCC TGG AAC TGT CCC CAT GGA AGG CCA ACC ATG AGA CAC ATC GCC AAC CTG GGT GTC ATT TCT
861 Gln Asn
2605 CAG AAC TGA CCG TAG TCA CTG TAT GGA ATA ATT GGT TTT ATC GCA GAT TTT TAT GTT TTG AAA GAC AGA GTC TTC ACT AAC CTT
2689 TTT TGT TTT AAA ATG AAA CCT GCT ACT TAA AAA AAA TAC ACA TCA CAC CCA TTT AAA AGT GAT CTT GAG AAC CTT TTC AAA CC ►

C		
yPMS1	mfhhienllietekrcqkqegryipvkylfsmtdhfhINDIDVHRITSGQVITDITAVKELVDNSIDAMNCEIIFKD	80
hPMS1	-----MKOLPAATVRLSSSOITTSVSVVKELIENSIDAGATSVVKLEN	46
hPMS2	meraessstepaka-----IKPIIDRKSVHQICSGQVVLSTIAVKELVENSIDAGATNIDLKLD	60
yPMS1	YGLSETECSNDGSGIDPSNYEFLALAHYTSKIAKQDVAKVQNLGFRGEALSSLCIAKLSVHTTSPPT-ADKLEYDMV	159
hPMS1	YGPDKIEVRDNGEGIKAVDAPVMAMRYTTSKINSHEDELENTYGFGEALSGICIAEVLITRTAADNFSTQVVLGS	126
hPMS2	YGVDLLEVSNDGCGVEENFEGITLPHHTSKIQEFADLTQVETFGFRGEALSSLCALSDVITSTCHASARVGTRLMFIHN	140
yPMS1	GHTTSKTTTSRNKCTTVLVSLPHNLPRVQKBFKSTFgrqftkcltviggyalinaaikfsvwnitpkjkknlilstmrn	239
hPMS1	CHILSQKPSHLGQCTTVTALRLKKNLPVRKQFYSTAKckckdeikkigdlmsfgilkpdlrivfvhnkviwqksrvsdh	206
hPMS2	CKTIQKTPYPRPRCTTVSVQQLRSTLPVRHKPFQRNikkeyakmvqvlhaycisagirvscntnqlgqkkrqpvvtcggs	220
yPMS1	ssmrkhlssvfgagggrgleevdlvidlnpfknrmkgkytdpdfldldykirvkgysisqnsfgcgrNSKDFCFIYVAKR	319
hPMS1	kmalmsevlgtavmnnnesfgyhseesqylsgflpkcdadhsfsl-----SIPERSFIFIMSR	265
hPMS2	psikenilgsvfggkqlqslipfvqlppsdsvceeyglscsdalhnlfyisgfisqcthgvr-----SNDRCFFIINRF	295
yPMS1	EVEYSTLRCCEVNYKTFnnv-----FAVFLNLEPMSLIDVNVTPDKRVILLHNERAVIDIFKTLSDYNNrnelalp	395
hPMS1	PVHQKDIKLRHHNLKcklestrlyPVFFLKTDVPTADVNLTPDKRSQVLLONKESVITALENMTTCGplpstns	345
hPMS2	PCDPAKVCILVNEVHMynrhz-----YFVVLNISVDEQCVINVTDPKROILLQEEKLLIAVLKTSIGMdsdvknln	371
yPMS1	krmcseqgaqqrkrktevfdirstthesdnenyhtarsesnqsnhahfnstgtvidksngteltsvmdgnyntvdvig	475
hPMS1	yennktvsaadivlsktaetdvlfnkvessgkynsvntsvipfndmndesgkntdcclnhqisigdfgyghcssei	425
hPMS2	vsqgplldvegnlikmhaadlkpmvkekdgqpslirtgeekkdvsisrlreaflrhttenkphspktpeprsrplgqkr	451
yPMS1	secevsdsssvldegnsstptkkipsiktddsqnlslnlfnnsfnpefnitspdksrlekvveepvyfididgkfkqek	555
hPMS1	snidkntknafqdismsnvswensqteysktcfissvkhqtsengnkdhidesgeneeaglnseisadewsrnlik	505
hPMS2	gmssstsgaisdkgvlrpqkeavssshgspdptdraevkdsgghstsvdsegfsipdtgshcsseyaasspddrgsge	531
yPMS1	avlsadglvfdvnechehtndcchqerrgstdeqddadsiaaeiepvveinvtpknrsksiskdnysrslsdglthr	635
hPMS1	nsvgsniepvkilypekslpckvsnnypipegmnlneascnkksnvidnksqkvtaydlslsnrvikpmsasalvfgdh	585
hPMS2	hvdskkapetddsfsvdchsnqedtgckfrvlpqptlatpntkrfkkeeilssadicqlvntqdmssasqdvavki	611
yPMS1	kfedeilevnlstknfkeiskngkqmsiiskrkseageniiknkdeledfeggekytlitvskndfkkmvvgqfnlgf	715
hPMS1	rpqfliennpksledatlqieelwktlseeeklyeekatkdlerynsqmkraieqesqmslkdgrrkkikptsawnlaqk	665
hPMS2	nkkvvpalfsmeslakrikqlhheaqqegeqnykrfrakicpggenqaadedlrkeisktmfaemeiigqfnlgfiitkl	691
yPMS1	iiivtrkvdnksdlfivdghasdekynfetlqavtvfksqkllipqvelsvidelvvldnlpvfekngfkldideeeefg	795
hPMS1	hkiktslsnqpkldellsgqiekrssqnikmvqipfsmnlknfkfkqknkvdeekdepclihnlrfpdawmtsktevm	745
hPMS2	nedifivdghatdekynfemlqghtvlgqgrliapqtlntatneavlienleifrkngdfvidenapvteraklislp	771
yPMS1	srvkllslptskqtlfdldgfnelihilhedgglrrdni-----	834
hPMS1	llnpyrvveallfkrllenhklpaeplepimlteslfnqshyldvlykmtaddqrgsgstylsdprltangfkiklpgg	825
hPMS2	tsknwtfgpqdvdelifmlsdspgvmc-----	798
yPMS1	-----RCSKIRSMFAMRACRSTIMGKPLNKKTITRVVHNLs	871
hPMS1	vsitenyleiegmanclpfygvadlkeilnailnrnakeyvecpPRKVISYLEGDAVLRLSRLPMYLSKEDQDITITRMK	905
hPMS2	-----RPSRVKQMBASRACRKSVMGTALTSEMKKLTITMg	835
yPMS1	eldkpw--NCPHGFTTMRHLMERdrwssfskdyei	904
hPMS1	hqfgneikeVHGPFTHFHHTYtpet-----	932
hPMS2	emdhpw--NCPHGFTTMRHLMANlgvisqm-----	862

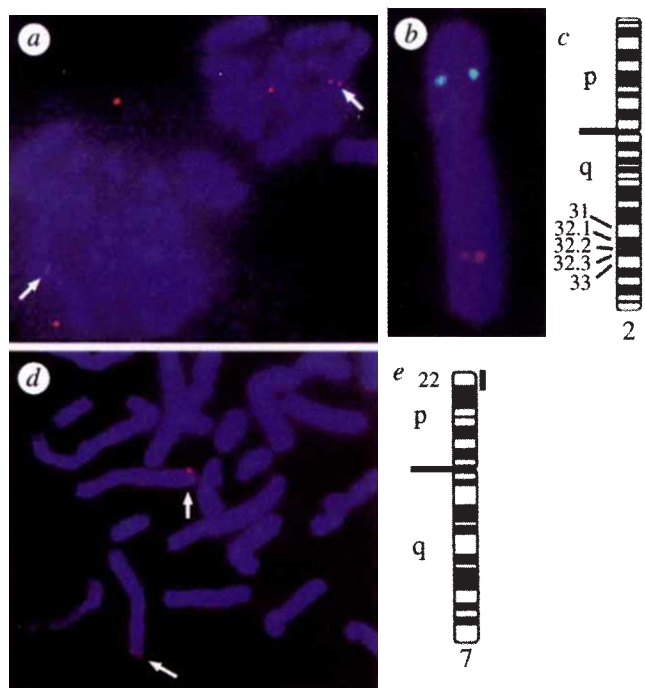


FIG. 2 Chromosomal localization. Metaphase spreads of normal human male chromosomes showing the in situ hybridization of *hPMS1* and *hPMS2*. **a**, Chromosome spread from a single cell showing hybridization of a *hPMS1* genomic clone to the distal q arm of both copies of chromosome 2 (arrows). A doublet signal on chromosome 2q was seen on more than 50 spreads from four separate experiments, while no doublets were detected on any other chromosomes. **b**, Simultaneous hybridization of *hPMS1* (green) and *hMSH2* (red) at chromosome 2p16. **c**, ISCN idiogram showing the predicted position (2q31-33) of *hPMS1* based on fractional length analysis of 14 individual chromosomes²⁸. **d**, Chromosome spreads probed with *hPMS2* which hybridizes near the telomere on the p arm of both copies of chromosome 7. A doublet signal on chromosome 7p was seen on more than 40 spreads from three separate experiments, while no doublets were detected on any other chromosomes. **e**, ISCN idiogram of *hPMS2* based on fractional length analysis of 6 individual chromosomes. This analysis indicated that the *hPMS2* gene was located within band 7p22, the most distal band on chromosome 7.

METHODS. A human genomic P1 library (Genome Systems, Inc.) was screened by PCR using primers selected from the cDNA sequence of *hPMS1* and *hPMS2*. Two P1 clones were isolated for *hPMS1* using primers 5'-AAGCTGCTCTGTAAAAGCG-3' and 5'-GCACGACATCCAAGAG-3' which yielded a 133 bp genomic product (nucleotide 58 to 190 of *hPMS1* cDNA). Three P1 clones were isolated for *hPMS2*, using primers 5'-CAACCATGAGACACATCGC-3' and 5'-AGGTTAGTGAAGACTCTGTC-3' which yielded a 121 bp genomic product (from nucleotide 2567 to 2687 of *hPMS2* cDNA.) FISH analyses were performed as described²⁸ except that nick-translated P1 clones were used as probe. Image alignment and fractional length analysis were done using the ISEE software package (Innovision Corp.), Durham, North Carolina.

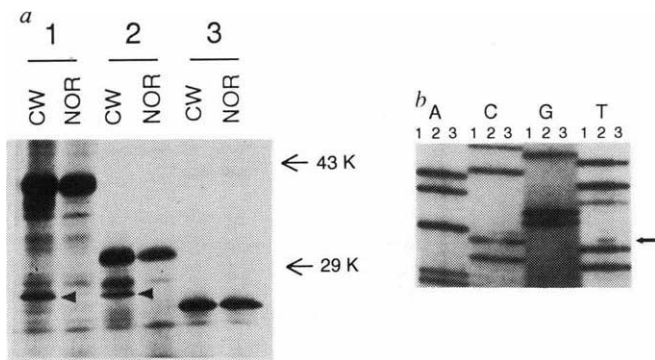


FIG. 3 Mutational Analysis of *hPMS1*. *a*, IVSP mapping of the transcriptional stop mutation in HNPCC patient C.W. Translation of codons 1 to 369 (lanes 1), codons 1 to 290 (lanes 2), and codons 1 to 214 (lanes 3). Lane CW was translated from the cDNA of patient C.W., while NOR was translated from the cDNA of a normal individual. The arrow heads indicate the truncated polypeptide due to the potential stop mutation. The arrows indicate relative molecular mass. *b*, Sequence analysis of C.W. indicates a C to T transition at codon 233 (indicated by the arrow). Lanes 1 and 3 are sequences derived from genomic DNA of control patients; Lane 2 is from genomic DNA of C.W.

METHODS. cDNAs from patient samples were generated from RNA of lymphoblastoid or tumour cells and used as templates in an IVSP assay as described²². The *hPMS1* gene was divided into three overlapping segments for the purpose of PCR. The primers for codons 1 to 500 were 5'-GGATCCTAATACGACTCACTATAGGGAGACCACCATGGAACAATTG-CCTGCGG-3' and 5'-CCTGCTCCACTCATCTGC-3'; for codons 270 to 755 were 5'-GGATCCTAATACGACTCACTATAGGGAGACCACCATGGAAGATATC-TTAAAGTTAATCCG-3' and 5'-GGCTTCTTCTACTCTATATGG-3'; and for codons 485 to 933 were 5'-GGATCCTAATACGACTCACTATAGGGAGAC-CACCATGGCAGGTCTTGAAACTCTTCG-3' and 5'-AAAACAAGTCAGTGATCCTC-3'. The 3' nested primers used for mapping the stop mutation in patient C.W. were: 5'-AAGCACATCTGTTTCTGCTG-3' codons 1 to 369; 5'-ACGAGTAGATTCTTTAGGC-3' codons 1 to 290; and 5'-CAGAACTGACATGAGAGCC-3' codons 15 to 214. The conditions for all amplifications were 35 cycles at 95 °C for 30 s, 52 °C to 62 °C for 60 to 120 s, and 70 °C for 60 to 120 s, in the buffer described previously²⁹. Intron-exon borders of *hPMS1* were determined by cycle-sequencing P1 clones using gamma-³²P end labelled primers and SequiTherm polymerase as described by the manufacturer. The primers used to amplify the *hPMS1* exon containing codons 195 to 233 were 5'-TTATTTGGCAGAAAGCA-GAG-3' and 5'-TTAAAAGACTAACCTCTTGCC-3', which produced a 215 bp product. The product was directly sequenced using the primer 5'-CTGCTGTTATGAACAATATGG-3'. PCR products were sequenced as described¹⁶.

protein in the absence of any full-length protein (Fig. 4*a*, lane GCx). In order to investigate the nature of this abnormality, the full-length RT-PCR products from G.C.'s normal fibroblasts and tumour xenograft were cloned and sequenced. Analysis of RNA from the normal fibroblasts revealed an altered transcript due to a 1230 bp in-frame deletion of codons 268 to 669 as well as the expected normal transcript. In contrast, analysis of the tumour xenograft identified a second transcript due to a somatic out-of-frame deletion which removed codons 301 to 381. The germline deletion of codons 268 to 669 was also observed in the tumour xenograft, but no normal transcript could be identified. It was likely that these altered transcripts were due to genomic deletions. To test this hypothesis, three sets of primers were used to amplify the *hPMS2* locus, one pair located 5' to the cDNA deletions (codons 233 to 257), one located 3' (codons 439 to 472), and one located in the commonly deleted region (codons 347 to 377) (Fig. 4*b*). Both the 5' and 3' regions of *hPMS2* could be amplified in all of the samples. The genomic segment containing codons 347 to 377 could be amplified from G.C.'s normal fibroblasts and from control patients but not from the tumour xenograft. These results indicated that the altered transcripts were the result of two intragenic deletions of *hPMS2*, one presumably germline and the other somatic. As

expected, the tumour from this patient exhibited microsatellite instability.

During the course of sequence analysis of the *hPMS2* cDNA, we noted a G to A transition in codon 20 which resulted in an arginine to glutamine change. This change was observed in 3 of 18 HNPCC patients and in 5 of 77 controls. We assume this change represents a polymorphism rather than a functional mutation.

The finding that human cells contain at least three *mutL* homologues suggests unique functions for each of these genes. Experimental evidence from *S. cerevisiae* has shown that at least two *mutL* related proteins are required for the major DNA mismatch repair pathway^{6,20}. The inactivation of either yeast *mutL* gene (*yMLH1* or *yPMS1*) results in a mutator phenotype^{6,20} and the encoded proteins interact in a complex²⁵. This view is supported in humans by the observation that inactivation of any one of three *mutL*-related genes can result in HNPCC, presumably due to mismatch repair deficiency. The identification of a

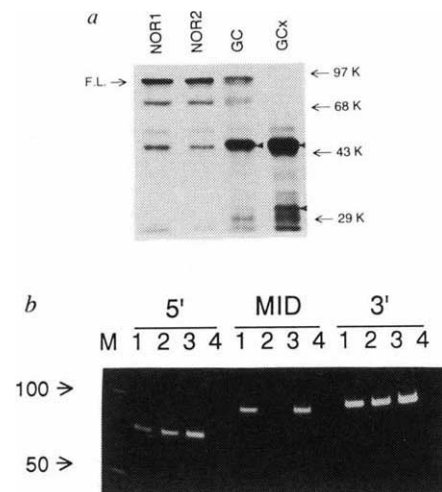


FIG. 4 Mutational analysis of *hPMS2*. *a*, IVSP analysis of *hPMS2* from patient G.C. Lane GC is from fibroblasts of individual G.C.; lane GCx is from the tumour xenograft of patient G.C.; lanes NOR1 and 2 are from normal control individuals. F.L. indicates full-length protein, and the arrowheads indicate the germline truncated polypeptide. A less intense low migrating polypeptide was also seen in all control samples. The arrows indicate relative molecular mass markers. *b*, PCR analysis of patient G.C. shows a deletion of both *hPMS2* alleles in tumour cells. Amplification was done using primers that amplify 5', 3', or within (MID) the region deleted in the cDNA. Lane 1, DNA derived from fibroblasts of patient G.C.; lane 2, DNA derived from tumour xenograft of patient G.C.; lane 3, DNA derived from a normal control patient; lane 4, reactions without DNA template. Arrows indicate size in base pairs. **METHODS.** IVSP analysis of *hPMS2* was done as described in Fig. 3 legend. In some cases, the PCR products were cloned into the T-tailed cloning vector PCR2000 (Invitrogen) and sequenced with T7 polymerase (United States Biochemical). The *hPMS2* cDNA was amplified as a full-length product or as two overlapping segments. The primers for full-length *hPMS2* were 5'-GGATCCTAATACGACTCACTATAGGGAGACCAC-CATGGAGCGAGCTGAGAGC-3' and 5'-AGGTTAGTGAAGACTCTGTC-3' (codons 1 to 863). For segment 1, the sense primer was the same as above and the antisense primer was 5'-CTGAGGTCTCAGCAGGC-3' (codons 1 to 472). Segment 2 primers were 5'-GGATCCTAATAC-GACTCACTATAGGGAGACCACCATGGTGTCCATTCCAGACTGCG-3' and 5'-AGGTTAGTGAAGACTCTGTC-3' (codons 415 to 863). The primers used to analyze the genomic deletion of *hPMS2* in patient G.C. were: For the 5' region amplification 5'-CAGAAGCAGTTGCAAAGCC-3' and 5'-AAACCGTACTCTTCACACAC-3' which produce a 74 bp product containing codons 233 to 257, for the deleted region primers 5'-GAG-GAAAAGCTTTTGTGGC-3' and 5'-CAGTGGCTGCTGACTGAC-3' which produce a 93 bp product containing the codons 347 to 377, and for the 3' region primers 5'-TCCAGAACCAAGAAGGAGC-3' and 5'-TGAGGTCTCAGCAGGC-3' which produce a 99 bp product containing the codons 439 to 472 of *hPMS2*.

somatic alteration of the remaining allele in a tumour from a patient with a germline *hPMS2* mutation supports the idea that inactivation of both alleles of a mismatch repair gene is required for tumour formation. □

Received 21 June; accepted 22 July 1994.

- Lynch, H. T. et al. *Gastroenterology* **104**, 1535–1549 (1993).
- Peinado, M. A., Malkhosyan, S., Velazquez, A. & Perucho, M. *Proc. natn. Acad. Sci. U.S.A.* **89**, 10065–10069 (1992).
- Ionov, Y., Peinado, A., Malkhosyan, S., Shibata, D. & Perucho, M. *Nature* **363**, 558–561 (1993).
- Thibodeau, S. N., Bren, G. & Schaid, D. *Science* **260**, 816–819 (1993).
- Aaltonen, L. A. et al. *Science* **260**, 812–816 (1993).
- Strand, M., Prolla, T. A., Liskay, R. M. & Petes, T. D. *Nature* **365**, 274–276 (1993).
- Parsons, R. P. et al. *Cell* **75**, 1227–1236 (1993).
- Umar, A. et al. *J. biol. chem.* **269**, 14367–14370 (1994).
- Peltomäki, P. et al. *Science* **260**, 810–812 (1993).
- Lindblom, A., Tannergard, P., Werelius, B. & Nordenskjöld, M. *Nature Genet.* **5**, 279–282 (1993).
- Nystrom-Lahti, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6054–6058 (1994).
- Fishel, R. et al. *Cell* **75**, 1027–1038 (1993).
- Leach, F. S. et al. *Cell* **75**, 1215–1225 (1993).
- Paiombo, F., Hughes, M. & Jiricny, J. *Nature* **367**, 417–418 (1994).
- Bronner, C. E. et al. *Nature* **368**, 258–261 (1994).
- Papadopoulos, N. et al. *Science* **263**, 1625–1629 (1994).
- Adams, M. D. et al. *Science* **252**, 1651–1656 (1991).
- Frohman, M. A., Dush, M. K. & Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8998–9002 (1988).
- Kozak, M. *Cell* **44**, 283–292 (1986).
- Prolla, T. A., Christie, D. M. & Liskay, R. M. *Molec. cell. Biol.* **14**, 407–415 (1994).
- Liu, B. et al. *Cancer Res.* (in the press).
- Powell, S. M. et al. *New Engl. J. Med.* **329**, 1982–1987 (1993).
- Van Der Luit, R. et al. *Genomics* **20**, 1–4 (1994). (Initials?)
- Dietz, H. C., et al. *Science* **259**, 680–683 (1993).
- Prolla, T. A. et al. *Science* (in the press).
- Kramer, W., Kramer, B., Williamson, M. S. & Fogel, S. J. *J. Bact.* **171**, 5339–5346 (1989).
- Schuler, G. D., Altschul, S. F. & Lipman, D. J. *Proteins Struct. Funct. Genet.* **9**, 180–190 (1991).
- Johnson, C. V. et al. *Meth. cell. Biol.* **35**, 73–99 (1991).
- Sidransky, D. et al. *Science* **252**, 706–709 (1991).

ACKNOWLEDGEMENTS. We would like to thank R. Wang, C. Fischer, V. Madike, and the sequencing staff from HGS for technical support. This work was supported by grants from the Clayton Fund, the Academy of Finland, the Finnish Cancer Societies, the Sigrid Juselius Foundation, the Folkhalsan Institute of Genetics, the Cooperative Human Tissue Network (NCI) and the NIH. N.C.N. and N.P. contributed equally to this work.

Crystal structure of an isoleucine-zipper trimer

Pehr B. Harbury^{*†}, Peter S. Kim[†] & Tom Alber[‡]

^{*} Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

[†] Howard Hughes Medical Institute, Whitehead Institute, Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA

[‡] Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA

SUBUNIT oligomerization in many proteins is mediated by short coiled-coil motifs^{1,2}. These motifs share a characteristic seven-amino-acid repeat containing hydrophobic residues at the first (a) and fourth (d) positions. Despite this common pattern, different sequences form two-, three- and four-stranded helical ropes. We have investigated the basis for oligomer choice by characterizing variants³ of the GCN4 leucine-zipper dimerization domain that adopt trimeric or tetrameric structures in response to mutations at the a and d positions. We now report the high-resolution X-ray crystal structure of an isoleucine-containing mutant that folds into a parallel three-stranded, α -helical coiled coil. In contrast to the dimer and tetramer structures^{3,4}, the interior packing of the trimer can accommodate β -branched residues in the most preferred rotamer at both hydrophobic positions. Compatibility of the shape of the core amino acids with the distinct packing spaces in the two-, three- and four-stranded conformations appears to determine the oligomerization state of the GCN4 leucine-zipper variants.

The GCN4-pII peptide, which differs from the wild-type GCN4 leucine zipper by isoleucine substitutions at four a and four d positions (Fig. 1a; Table 1), forms a parallel trimer in solution³. A similar isoleucine repeat occurs in the trimeric σ 1 protein of reovirus serotype 1 (refs 5, 6) in haemagglutinins from several strains of human parainfluenza virus^{7,8} and in trout axonemal dynein⁹. The trimeric coiled coils from influenza haemagglutinin^{10,11} and the yeast heat-shock factor¹² (HSF) resemble the GCN4-pII sequence at the d positions of the heptad repeat, which contain predominantly β -branched residues (IIIL in the five heptads of HSF).

The X-ray crystal structure of GCN4-pII at 1.8 Å resolution (Fig. 1) shows that three α -helical peptide monomers wrap in a gradual left-handed superhelix. The superhelix describes a cylinder that is ~24 Å wide and ~48 Å long. The isoleucine residues point into the core of the trimer, and cross-sectional layers containing isoleucine at the a positions alternate with layers containing isoleucine at the d positions (Fig. 1d). In seven of eight layers, the dihedral angles χ_1 and χ_2 of the isoleucine side chains are approximately –60,180, corresponding to the most abundant rotamer¹³. The isoleucine residues in the C-terminal a layer assume dihedral angles near +60,180. Residues at positions e and g pack against the isoleucines at d and a, respectively, to complete the hydrophobic core. Interhelical salt bridges form with high frequency between charged side chains at the g position of one heptad and the e position of a succeeding heptad (for

TABLE 1 Helix-helix interactions*

	Dimer	Trimer	Tetramer	Globular proteins ¹⁷
GCN4 peptide variant†	pI‡ and pL	pI	pL	
Residues at four a positions	V(N)	I	L	
Residues at four d positions	L	I	I	
Superhelix parameter				
Supercoil radius, R_0 (Å)	4.9	6.7	7.6	
Residues per supercoil turn, ω_0	100	118	139	
Supercoil pitch (Å)	148	175	205	
Radius of curvature (Å)	118	124	149	
Superhelix crossing angle (χ)	23.4°	26.8°	26.0°	
Position a orientation angle, ϕ	21.6°	20.4°	19.8°	
α-Helix parameter¹⁷				
Residues per α -helix turn, n	3.62	3.60	3.59	3.64 ± 0.18
Rise/residue, d (Å)	1.51	1.53	1.52	1.51 ± 0.12
α -Helix radius ($C\alpha$), R_1 (Å)	2.28	2.24	2.26	—
Pairwise helix-crossing angle, Ω	23.4°	23.2°	18.3°	$19^\circ \pm 24^\circ$
Pairwise interhelix distance, D (Å)	9.8	11.5	10.6	10.2 ± 2.0
Interhelical salt bridges				
g to e	3 of 6	8 of 9	5 of 12	
g to b and c to e	0 of 8	1 of 12	9 of 16	

* Superhelical characteristics were obtained by fitting the $C\alpha$ backbones to a supercoil parametrization suggested by Crick²² (Fig. 1a). The radius (R_0), frequency (ω_0) and pitch of the superhelix, and the radius (R_1) and phase (ϕ) of the α -helix were treated as variables. The frequency of the α -helix was fixed at 4π radians per 7 amino acids to preserve the periodicity of the heptad structure. α -helix parameters were calculated as defined in ref. 17. The properties of α -helices in globular proteins were taken from that reference and are expressed as the mean value ± 2 standard deviations. The values of D and Ω for globular proteins are averages for the 3–4 ridges-into-grooves packing class¹⁷. Consistent with earlier analyses^{3,22,23}, the superhelical radius and pitch, and the radius of helix curvature increase with the number of strands in the coiled coil.

† The leucine zipper variants were derived from the wild-type GCN4 coiled-coil peptide, GCN4-p1 (ref. 24), by collective mutation to the indicated residues at four a positions (dashed box in Fig. 1a) and four d positions (dashed oval in Fig. 1a)³.

‡ Geometric properties of the dimer conformation were obtained from the structure of GCN4-p1 (ref. 4) because the crystallographic analysis of GCN4-pIL has not been completed. The GCN4-p1 sequence contains four leucines at the varied d positions, and three valines and one asparagine at the varied a positions. Interestingly, replacement of the asparagine residue (Asn 16) by valine generates a peptide, GCN4-pVL, that forms a mixture of dimers and trimers³. This result indicates that the buried polar side chain of Asn 16 directs dimer formation.

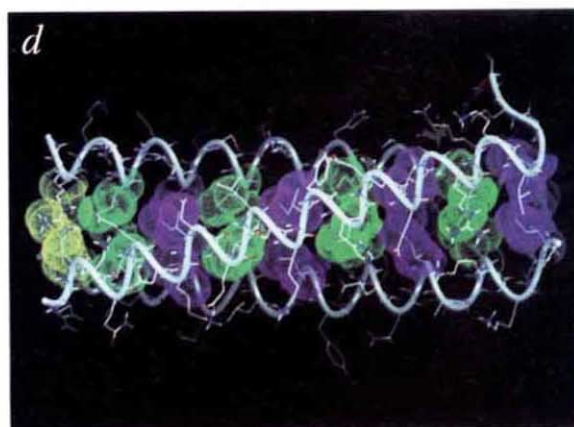
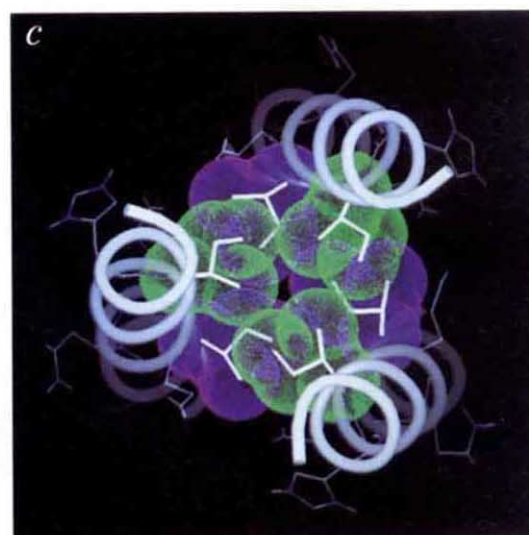
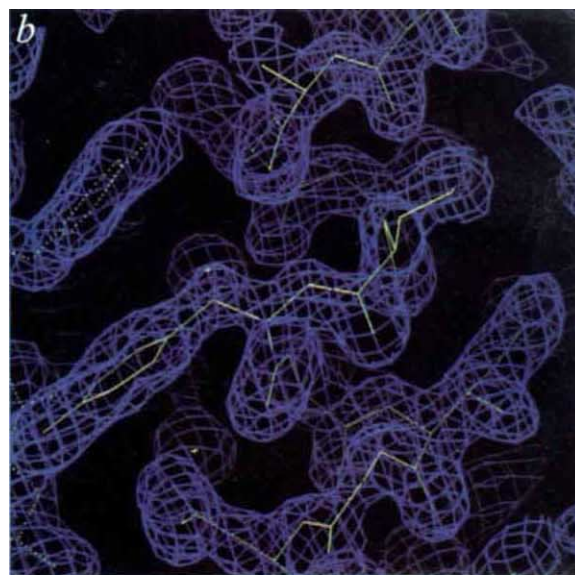
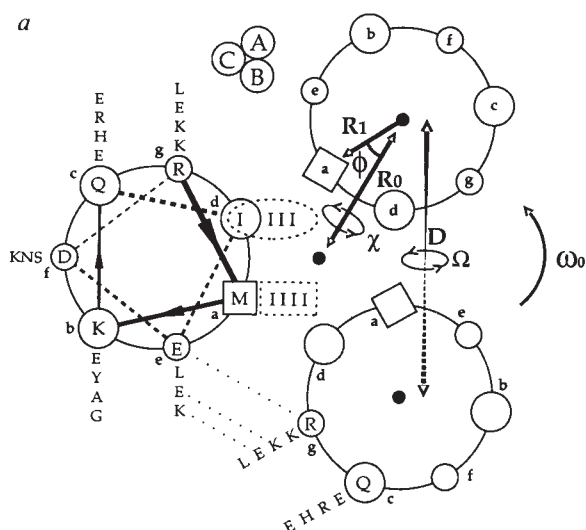


FIG. 1 Crystal structure of the GCN4-pII trimer. **a**, Helical wheel projection of residues 1 (Arg) to 32 (Glu) of the GCN4-pII sequence. View is from the N terminus, and residues in the first two helical turns are boxed or circled. Heptad positions are labelled **a–g**. The R_0 , ω_0 , α , ϕ and R_1 variables from a coiled-coil parameterization suggested by Crick^{3,22} and the D and Ω variables defined in ref. 17 are illustrated. **b**, A portion of the 6.0–1.8 Å resolution $2F_o - F_c$ electron-density map superimposed on the final model. A side view of residues Ser C14 to Asn C21 is shown. **c**, Axial view of the GCN4-pII trimer. The van der Waals surfaces are coloured purple for residues at **a** positions and green for residues at **d** positions. **d**, Side view of the trimer. The van der Waals surfaces of the N-terminal methionine layer are yellow.

METHODS. The GCN4-pII peptide, Ac-R MKQIEDK IEEILSK IYHIENE IAR IKKLIGER (where mutant **a** residues are underlined and mutant **d** residues double-underlined; Ac, acetyl), was crystallized by vapour diffusion from 100 mM sodium citrate, pH 4.8, 600 mM NaBr and 5% PEG 400. The crystals had $P2_1$ symmetry ($a = 35.32$ Å, $b = 39.67$ Å, $c = 36.97$ Å, $\beta = 98.38^\circ$) with a trimer in the asymmetric unit. Data to 1.8 Å resolution were collected using an *R*-Axis image plate detector (7,052 unique reflections, $R_{\text{merge}} = 0.045$). An idealized model was built²⁵ using an empirical method for calculation of coiled-coil structures based on an algebraic parameterization²². The model was used in a combined rotation–translation search (with 6–2.9 Å resolution data $> 2\sigma$) along a non-crystallographic 3-fold symmetry axis (located with GLRF²⁶) using the program XPLOR²⁷. The coordinates corresponding to the peak solution (initial $R = 52\%$) were annealed and rebuilt into simulated-annealing omit maps²⁸. Water molecules were added and the coordinates and individual B values were refined using TNT²⁹ and FRODO³⁰. The final model, which includes 96 amino acids of the trimer and 48 water molecules, has a crystallographic R -factor of 0.175 (6–1.8 Å data) with root-mean-square deviations from ideal bond lengths and bond angles of 0.013 Å and 2.4° respectively. All main-chain dihedral angles fall within allowed regions of the Ramachandran plot, and most side chains assume well populated rotamer conformations. The coordinates of the search model and the refined structure differ by a root-mean-square deviation of 0.46 Å for main-chain and core side-chain atoms. These results demonstrate the feasibility of using a predicted structure to initiate crystallographic analysis. The coordinates have been deposited in the Brookhaven Protein Data Bank.

example, Glu B22–Lys C27; Table 1 and Fig. 1a). This observation supports proposals that heterotrimer formation could be directed by electrostatic attraction¹⁴ or repulsion^{15,16} between these residues.

Although the pitch, radius and residues per turn of the superhelices in the trimer (GCN4-pII), dimer (GCN4-pI) and

tetramer (GCN4-pLI) differ greatly, the individual α -helices are virtually identical on a local scale (Table 1). Moreover, pairs of adjacent helices in the three structures have similar separations (D) and crossing angles (Ω). This conserved pairwise arrangement of helices corresponds to the optimal geometry for one class of helix association in globular proteins¹⁷ (Table 1).

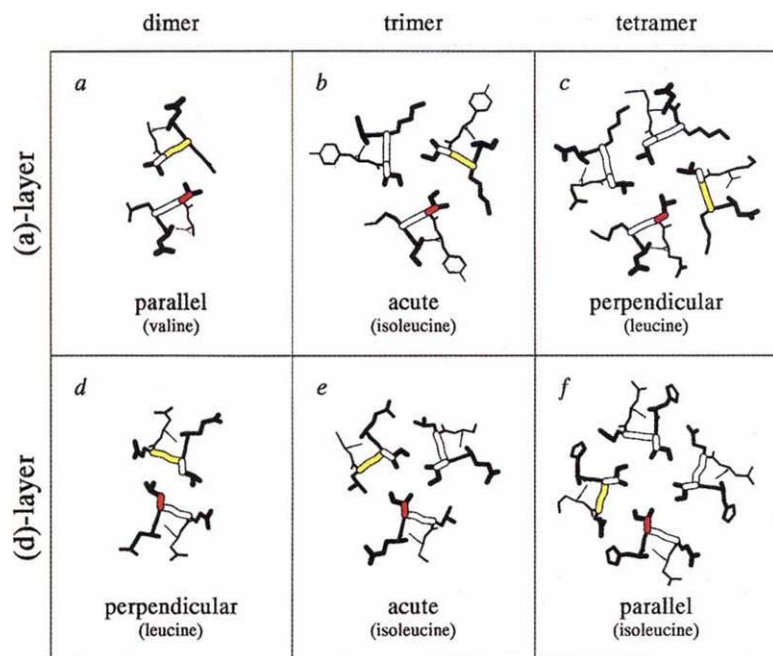


FIG. 2 Packing differences in parallel coiled coils^{3,18}. Helix cross-sectional layers centred on the **a** positions (containing the residues at positions **f**, **g**, **a** and **b**) and on the **d** positions (containing the residues at positions **c**, **d**, **e** and **f**) are depicted³¹. Knobs formed by the side chains of one helix fit into holes formed by the spaces between side chains on a neighbouring helix. Looking from the N terminus down the superhelix axis, each knob at an **a** position (red) packs into a hole formed between the **g** and **a** residues (yellow) and two **d** residues in adjacent layers (not shown) of the anti-clockwise-related monomer. Similarly, each knob at a **d** position (red) packs into a hole formed between the **d** and **e** residues (yellow) and two **a** residues in adjacent layers (not shown) of the clockwise related monomer. *a*, Position **a** of the dimer: the $Ca-C\beta$ bond of each valine knob (red) is oriented parallel to the $Ca-Ca$ vector at the base of the hole into which it projects (yellow). *b*, Position **a** of the trimer; the $Ca-C\beta$ bond of each isoleucine knob (red) forms an acute angle ($\sim 50^\circ$) with the $Ca-Ca$ vector at the base of the recipient hole (yellow). *c*, Position **a** of the tetramer; the $Ca-C\beta$ bond of each leucine knob (red) is perpendicular to the $Ca-Ca$ vector at the base of the apposed hole (yellow). *d*, Perpendicular packing at position **d** of the dimer. *e*, Acute packing at position **d** of the trimer. *f*, Parallel packing at position **d** of the tetramer.

The GCN4-pII trimer shows 'knobs-into-holes' packing, in which the side chains (knobs) at positions **a** and **d** fit into the spaces (holes) between four residues on a neighbouring helix¹⁸. In both the **a** and **d** layers, the $Ca-C\beta$ bond of the knob side chain makes an acute angle with the $Ca-Ca$ vector defining the base of the apposed hole (Fig. 2*b, e*). This shared geometry results in a similar placement of atoms around the side chains at positions **a** and **d** (Fig. 3*b*). However, the acute angles in the two types of layers differ by $\sim 25^\circ$ (Fig. 3*b*). In addition, the third helix of the trimer sits to the C-terminal side of the hole in the **a** layer (anti-clockwise from hole in Fig. 2*b*) but to the N-terminal side of the hole in the **d** layer (clockwise from hole in Fig. 2*e*). Thus, for example, a threonine knob residue with a $(-)\chi_1$ dihedral angle would direct its hydroxyl group into the hydrophobic interface at an **a** position, but towards solvent at a **d** position.

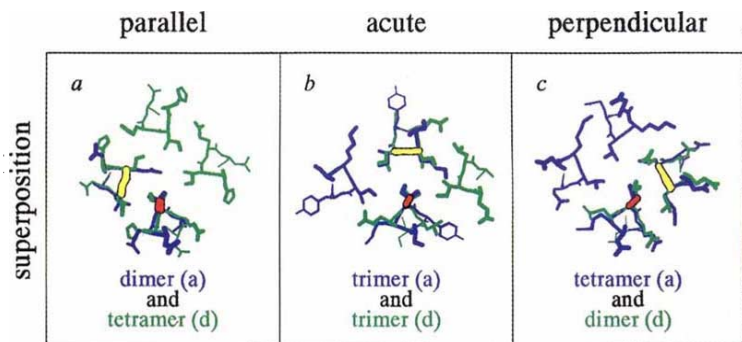
The interior packing of the trimer differs from that seen in the dimeric and tetrameric leucine-zipper variants. In the GCN4-pI dimer, the $Ca-C\beta$ bond of the knob side chain at position **a** points out of the interface (parallel packing; Fig. 2*a*), and the $Ca-C\beta$ bond of the knob side chain at position **d** points directly at the adjacent helix (perpendicular packing; Fig. 2*d*). This pattern is reversed in the GCN4-pLI tetramer (Figs 2*c*, and 3*a, c*), which shows perpendicular packing in the **a** layers and parallel packing in the **d** layers. Thus the packing geometries in the dimer, trimer and tetramer are distinct from one another at both hydrophobic positions (Fig. 2).

The three modes of knobs-into-holes packing exhibit different preferences for specific amino acids at the knob position. In dimeric coiled-coil sequences, the **a** positions (parallel packing) are enriched for the β -branched residues valine and isoleucine, whereas the **d** positions (perpendicular packing) are sharply depleted of β -branched residues and enriched for leucine^{19,20}. In trimeric coiled-coil sequences, however, leucine and the β -branched amino acids are approximately equally represented and distributed evenly between the **a** and **d** positions²¹ (consistent with the geometric similarity of the packing at **a** and **d** in the trimer; Fig. 3*b*). Thus parallel geometry favours β -branched residues, perpendicular geometry disfavors β -branched residues and favours leucine, and acute geometry shows little preference.

The oligomerization states of the GCN4 leucine-zipper mutants are apparently dictated by matching the core residues, determined by sequence, with appropriate packing geometry, determined by structure. The dimer of GCN4-pIL and the tetramer of GCN4-pLI, for example, place the leucine residues in perpendicular geometry and the isoleucine residues in parallel geometry. By contrast, a dimer or tetramer of GCN4-pII would mismatch four of the eight buried isoleucines with perpendicular geometry. The trimer conformation, however, places all the isoleucines in acute geometry, thereby accommodating β -branched amino acids at both the **a** and **d** positions.

The packing properties of the β -branched residues isoleucine and valine are not identical. For example, the peptide

FIG. 3 Superposition³¹ of parallel, perpendicular and acute knobs-into-holes packing. *a*, Parallel packing in the dimer **a** layer (green) and the tetramer **d** layer (blue). *b*, Acute packing in the trimer **a** layer (green) and the trimer **d** layer (blue). *c*, Perpendicular packing in the tetramer **a** layer (green) and the dimer **d** layer (blue).



GCN4-pVL forms both dimers and trimers and the peptide GCN4-pLV is trimeric³. Comparison of the oligomerization properties of GCN4-pVL and GCN4-pIL (dimer), or GCN4-pLV and GCN4-pLI (tetramer), suggests that isoleucine exhibits a stronger preference than valine for parallel over acute geometry. Thus, the GCN4 variant coiled coils provide a striking example of how side-chain shape and packing can dramatically affect the global fold of a protein. □

Received 27 April; accepted 8 July 1994.

1. Cohen, C. & Parry, D. A. D. *Proteins* **7**, 1–15 (1990).
2. Cohen, C. & Parry, D. A. D. *Science* **263**, 488–489 (1994).
3. Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. *Science* **262**, 1401–1407 (1993).
4. O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. *Science* **254**, 539–544 (1991).
5. Duncan, R., Horne, D., Cashdollar, L. W., Joklik, W. K. & Lee, P. W. K. *Virology* **174**, 399–409 (1990).
6. Strong, J. E., Leone, G., Duncan, R., Sharma, R. K. & Lee, P. W. K. *Virology* **184**, 23–32 (1991).
7. Elango, N., Coligan, J. E., Jambou, R. C. & Venkatesan, S. *J. Virol.* **57**, 481–489 (1986).
8. Coelingh, K. L. V. W., Winter, C. C. & Murphy, B. R. *Virology* **162**, 137–143 (1988).
9. Garber, A. T., Retief, J. D. & Dixon, G. H. *EMBO J.* **8**, 1727–1734 (1989).
10. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366–373 (1981).
11. Carr, C. M. & Kim, P. S. *Cell* **73**, 823–832 (1993).
12. Sorger, P. K. & Nelson, H. C. *Cell* **59**, 807–813 (1989).
13. Ponder, J. W. & Richards, F. M. *J. molec. Biol.* **193**, 775–791 (1987).
14. Beck, K., Dixon, T. W., Engel, J. & Parry, D. A. *J. molec. Biol.* **231**, 311–323 (1993).
15. O'Shea, E. K., Rutkowski, R. & Kim, P. S. *Cell* **68**, 699–708 (1992).
16. O'Shea, E. K., Lumb, K. J. & Kim, P. S. *Curr. Biol.* **3**, 658–667 (1993).
17. Chothia, C., Levitt, M. & Richardson, D. J. *J. molec. Biol.* **145**, 215–250 (1981).
18. Crick, F. H. C. *Acta crystallogr.* **6**, 689–697 (1953).
19. Hu, J. C. & Sauer, R. T. in *Progress in Nucleic Acids and Molecular Biology* (eds Eckstein, F. & Lilley, D. M. J.) 82–101 (Springer, Berlin, 1992).
20. Conway, J. F. & Parry, D. A. D. *Int. J. biol. Macromol.* **12**, 328–334 (1990).
21. Conway, J. F. & Parry, D. A. D. *Int. J. biol. Macromol.* **13**, 14–16 (1991).
22. Crick, F. H. C. *Acta crystallogr.* **6**, 685–689 (1953).
23. Seo, J. & Cohen, C. *Proteins* **15**, 223–234 (1993).
24. O'Shea, E. K., Rutkowski, R. & Kim, P. S. *Science* **243**, 538–542 (1989).
25. Harbury, P. H. thesis, Harvard Univ. (1994).
26. Tong, L. & Rossmann, M. G. *Acta crystallogr.* **A46**, 783–792 (1990).
27. Brunger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
28. Brunger, A. T. *J. molec. Biol.* **203**, 803–816 (1988).
29. Tronrud, D. E., TenEyck, L. F. & Matthews, B. W. *Acta crystallogr.* **A43**, 489–501 (1987).
30. Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
31. Kraulis, P. J. *J. appl. Cryst.* **24**, 946–950 (1991).

ACKNOWLEDGEMENTS. We thank M. Milhollen for peptide synthesis. This work was supported by grants from the National Institutes of Health (to P.S.K. and T.A.) and the Lucille P. Markey Charitable Trust (to T.A.).

CORRECTION

Flattening of the sea-floor depth–age curve as a response to asthenospheric flow

Jason Phipps Morgan & Walter H. F. Smith

Nature **359**, 524–527 (1992)

WE are grateful to Carol and Seth Stein¹ for alerting us to errors in our discussion of the results shown in Fig. 2b of our Letter. In the calculations shown in this Figure, the asthenospheric viscosity beneath the South Atlantic lithosphere was 4.2×10^{19} Pa s and the African plate was assumed to be stationary, instead of migrating eastward at 10 mm yr^{-1} as stated in the Letter. In addition, the thermal subsidence rate and ridge-crest depth, which were not mentioned, were $280 \text{ m Myr}^{-1/2}$ and $2,200 \text{ m}$ respectively. In contrast, the values used for the Pacific plate were, as stated, $300 \text{ m Myr}^{-1/2}$ and $2,644 \text{ m}$. Hence the fit to the Pacific and South Atlantic plates results from assuming subsidence rates differing by 7% and viscosities differing by a factor of 21. □

1. Stein, C. & Stein, S. *Geophys. Res. Lett.* **21**, 709–712 (1994).

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without display items, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are processed from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review Articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Progress articles review particularly topical developments for a nonspecialist readership. They do not exceed 4 pages in length. Suggestions may be made to the Reviews Coordinator in the form of a brief synopsis.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research.

News and Views articles inform nonspecialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required, for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers, are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books includes publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary Information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. A telephone and fax number should be included. Manuscripts or proofs sent by air courier to London should be declared as manuscripts and 'value \$5' to prevent the imposition of import duty and value-added tax.

Shorting out the cellular battery

A low-resolution structure of the bacterial toxin colicin Ia provides a model for its lethal membrane association and an explanation for the anomalous behaviour of its C-terminal peptide fragment.

POSSIBLY the most significant single feature that distinguishes living cells from dead ones is their ability to establish and maintain an electrochemical potential across their plasma membranes. By pumping specific ions (in particular Na^+ and K^+) across the cell membrane, a potential of -20 to -200 mV is produced, depending on organism and cell type. Without this potential the cell dies, a fact exploited by a number of toxins which kill their target cells by depleting this potential in an action exactly analogous to shorting out an electric battery. To achieve this, the toxin punches holes in the cell membrane causing ions to rush across and the potential to collapse.

Among the toxins that act in this fashion are a group of proteins produced by *Escherichia coli* known as the pore-forming colicins. Colicins are encoded on a bacterial plasmid which also conveys resistance to colicin attack. A small proportion of plasmid-bearing cells will be producing colicin at any one time, destroying any bacteria not bearing the colicin plasmid and so clearing the way for those that do. In order to penetrate the complex bacterial cell wall these toxins undergo a series of structural changes, but until now little has been known of their structure and less still of their mechanism of action. Now with the first structure of an intact colicin, colicin Ia (ref. 1), albeit at relatively low resolution ($4\text{\AA}$), a unified picture of their bacteriolytic activity is beginning to emerge^{1,2}.

Ghosh and colleagues¹ have succeeded in crystallizing the intact soluble form of colicin Ia and have calculated a $4\text{\AA}$ electron density map. Although at this resolution it is not possible to follow the amino-acid chain through the protein, regular pillars of electron density indicate the positions of α -helices. The locations of a few individual amino-acid residues could also be assigned, giving confidence in the accuracy of the electron density map. Colicin Ia is clearly seen to be composed of several α -helices, linked by

irregular regions which cannot be seen at this resolution. The structure is divided into three domains arranged in a rough 'Y' shape, $95\text{\AA}$ long and $75\text{\AA}$ across the two arms. By a synthesis of previous biochemical data and an earlier structure of a fragment from a related protein, colicin A (ref. 3), it can be shown that the two arms are formed by the carboxy and central two-thirds of the molecule while the 'stem' comprises the amino terminus. A separate function has been localized to each of these three domains.

The initial step in bacterial cell attack is attachment to the outer bacterial membrane by binding to receptors on the target cell's surface. This is a process common to many cellular toxins and the specific receptors are frequently carbohydrate moieties: see for example the structure of pertussis toxin bound to an undecasaccharide in this month's *Nature Structural Biology*⁴. For colicin Ia this function is mapped to the central region of the protein which forms one of the arms of the 'Y' (magenta in the figure).

After binding to its receptor, the toxin passes across the periplasmic space to reach the inner bacterial membrane. This translocation requires the N-terminal portion of the molecule which forms the stem of the 'Y' and is made up of three α -helices, the longest of which is about $70\text{\AA}$. This means that colicin Ia may attach to the bacterial outer membrane by this N-terminal bridging domain while the arms of the 'Y' stretch out to touch the inner cell membrane.

Finally, the carboxy-terminal third of the molecule (yellow and orange in the figure) inserts into the inner cell membrane, forming a nonspecific ion channel. This is a two-step reaction: two hydrophobic α -helices insert into the membrane as a β -hairpin then, in the presence of an electrical potential, more helices are driven into the membrane to form a voltage-gated pore. This pore remains open until the membrane potential rises to around $+90$ mV, a complete reversal of the normal situation.

C-terminal fragments of a number of colicins, including colicin Ia, can form ion channels in lipid membranes *in vitro*, but their physico-chemical properties (voltage dependence, open times and so on) are different from those of the intact protein. It is fragments of this kind whose structures have previously been determined^{3,5}. These are composed of a



three-layered bundle of ten α -helices formed around the two hydrophobic helices that initially insert into the membrane. In the complete molecule this fragment is seen as part of one arm of the 'Y' (yellow), but there are also another eight α -helices making up the rest of the 'arm' (orange). This subdomain may well be responsible for the differences in activity between C-terminal pore-forming fragments and the intact colicins. Our models of membrane insertion, derived primarily from isolated pore-forming fragments, will now need to be recast.

As so often with solutions of protein structures, more questions have been raised than answered, answers that must come from a combination of different techniques. This structure also highlights the difficulty of interpreting data from protein fragments without knowing how such fragments fit into the whole. Now with our clearer picture of colicin Ia, and therefore of other homologous colicins, studies of its attachment, translocation and membrane insertion can proceed apace. This work on what is a relatively simple toxin can now be used to model the activity of more complex toxins, including the causative agents of diphtheria and cholera.

Christopher Surridge

Christopher Surridge is Assistant Editor of Nature Structural Biology.

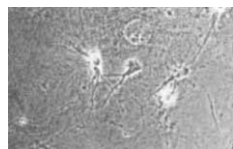
Also in this month's *Nature Structural Biology*: detailed structures of a TATA-box-binding protein with and without bound DNA; an isolated β -hairpin free in solution; identifying the DNA-binding site of heat-shock transcription factor; and the structure of a pertussis toxin-sugar complex.

1. Ghosh, P., Mel, S.F. & Stroud, R.M. *Nature struct. Biol.* **1**, 597-604 (1994).
2. Sansom, M., *Nature struct. Biol.* **1**, 563-567 (1994).
3. Parker, M.W., Postma, J.P.M., Pattus, F., Tucker, A.D. & Tsermoglou, D. *J. molec. Biol.* **224**, 639-657 (1992).
4. Stein, P.E. et al. *Nature struct. Biol.* **1**, 591-596 (1994).
5. Wormald, M.R., Merrill, A.R., Cramer, W.A. & Williams, R.J.P. *Eur. J. Biochem.* **191**, 155-161 (1990).

Going for growth

Cell and tissue culture aids featured this week include postmitotic human neuronal cells, viability test kits for yeast, fungi and bacteria, cultureware, bioreactors, platers, shakers, stirrers and workstations.

STRATAGENE Cloning Systems and Layton BioScience have signed an agreement that gives Stratagene the exclusive right to market Layton's hNT **human neuronal cells**



Fully differentiated hNT human neurons.

for research use in the United States, Canada and Europe (*Reader Service No. 100*). Developed by Virginia M. -Y Lee and John Q. Trojanowski

at the University of Pennsylvania, which licensed the cells exclusively to Layton, the postmitotic human neuronal cells are intended for use in a number of applications—studies of biological function, growth and differentiation; drug screening; *in vitro* neurotoxicity testing; and, ultimately, transplantation to replace damaged brain cells. The cells may also offer new avenues of research into Alzheimer's disease and AIDS dementia complex.

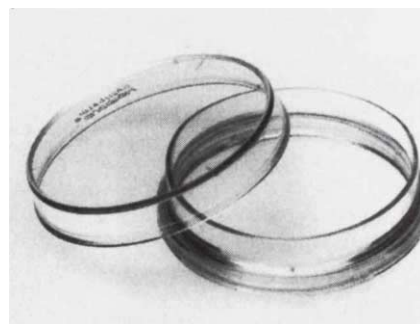
SM ID is bioMérieux's ready-to-use **plated medium for the detection of *Salmonella* strains** (*Reader Service No. 101*). This new selective medium contains the two substrates: glucuronate and β -galactosidase. Positive identification of *Salmonella* is based on lack of β -galactosidase activity coupled with the metabolism of glucuronate, which causes the formation of pink colonies after incubation for 18–24 h at 37 °C. SM ID is supplied in packs of twenty 90-mm plates.

Molecular Probes has developed an assay for detecting yeast and fungi and assessing their viability—the **Live/Dead FungoLight one-step viability test kit for yeast and fungi** (*Reader Service No. 102*). The kit includes the dye FUN-1 and the counterstain Calcofluor White M2R. It can be used to discriminate yeast in mixed-cell populations as bacterial contaminants of yeast cultures are said to stain uniformly fluorescent green. Viable yeast cells can also be detected in biologically complex mixtures such as blood. Also available, an alternative to more laborious conventional bacterial viability tests—the **Live/Dead BacLight one-step bacterial viability kit**. This assay can be used to distinguish live bacteria with intact plasma membranes from dead bacteria with compromised membranes. Live bacteria will fluoresce a brilliant green colour whereas dead cells will show up as an intense red colour. The kit can be used to determine the viability of bacteria in mixed-

cell populations and in samples that contain other cell types, including eukaryotic cells. The proportion of live and dead bacteria in a suspension can also be determined quantitatively using a fluorometer, fluorescence microplate reader or flow cytometer.

■ Culturing aids

Petriperm is the new disposable, sterile **tissue culture dish with a gas-permeable hydrophilic or hydrophobic growth surface** now available from Specialized Cultured Technologies (*Reader Service No. 103*). The 20 cm² growth surface is made of a 25- μ m thick Teflon foil, said to be as clear as glass and nontoxic for cells. Most primary cell types can be cultured on either the hydrophilic (adherent) or hydrophobic (non-adherent) growth surface under different oxygen tensions. Applications include whole



embryo and organ culture, embryonic stem cell and primordial germ cell cultures, long-term cultures on adherent cell layers, stem cell expansion, macrophage and lymphocyte cultures, cell cloning and studies on cell-matrix interactions. The company says that fixation and staining of cells for immunocytochemistry or histochemistry can be performed directly in the dish. In addition, most light microscopy techniques can be performed, for example, bright and dark field, oil immersion, phase and reflection contrast, polarization, differential interference contrast and fluorescence, as well as electron microscopy.

Becton Dickinson/Collaborative Biomedical Products has introduced a new line of **plasticware manufactured for *in vitro* fertilization** (*Reader Service No. 104*). Falcon IVF Plasticware is made from crystal-grade polystyrene. All products are said to be sterile, tested for the absence of mouse embryotoxicity, and certified to be non-pyrogenic. The new product line includes

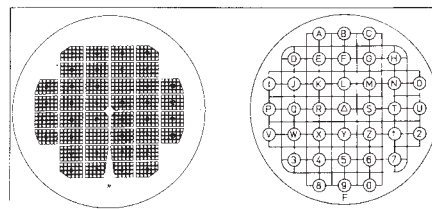
60-mm dishes, available with a tissue culture-treated hydrophilic surface (TC dish), or a non-treated hydrophobic surface (Petri dish), a 1-well organ culture dish, and a 4-well tissue culture plate. Suggested uses for the plasticware include the *in vitro* fertilization of human oocytes and the *in vitro* manipulation of animal oocytes and embryos (that is, for producing transgenic animals).

The BioCoat **leucocyte traffic environment** is an *in vitro* system for the study of leucocyte adhesion to, and migration across, endothelium, which is now available from Stratech Scientific (*Reader Service No. 105*). The company says that with this system researchers can establish confluent endothelial monolayers with barrier function within 2 to 3 days. Also offered cell growth environments for growing normal keratinocytes and endothelial cells, and a system containing the materials necessary for the retention of the differentiated phenotype of primary hepatocytes in serum-free conditions.

Costar has developed a range of polycarbonate- and collagen-coated permeable supports for applications ranging from chemotaxis to transport studies. As a complement to its current range of **permeable supports**, Costar has introduced Transwell Clear supports, where the optimized pore density of the membrane used makes it particularly suitable for transport studies, chemotaxis, co-culture and microbial pathogenesis (*Reader Service No. 106*).

■ Gadgets

Merck has introduced **CELLocate microgrid coverslips** from Eppendorf designed specifically for cell technology (*Reader Service No. 107*). The glass microgrid coverslip is placed on the base of a cell culture dish and enables individual cells or cell groups to be relocated and the population of the dish to be divided into distinct subpopulations. In this way, a direct comparison of cells undergoing different treatments within the same basic set of conditions can be made. In



Microgrid coverslips.

PRODUCT REVIEW

addition, CELLocate can be used for measuring cell growth and cytotoxicity, determining cell division indices and estimating cell numbers. CELLocate is available in packs of either 80 or 200 sterile or nonsterile coverslips and with a choice of two grid sizes: 55 μm and 175 μm . Suitable for use in fluorescence and electron microscopy.

Temp Guard is a **battery-operated, temperature monitoring device** available from Manostat that sounds an alarm when the temperature of the incubator is out of range (*Reader Service No. 108*). And, even



Temperature alarm.

if the variance in temperature is corrected, an LED stays lit to record that the temperature had changed, and to alert users to possible damage to the contents of the incubator. Intended for cell culture and any laboratory

where temperature-sensitive supplies are stored, Temp Guard can be mounted on or near any incubator with either self-adhesive tape or screw mountings.

■ Culture systems

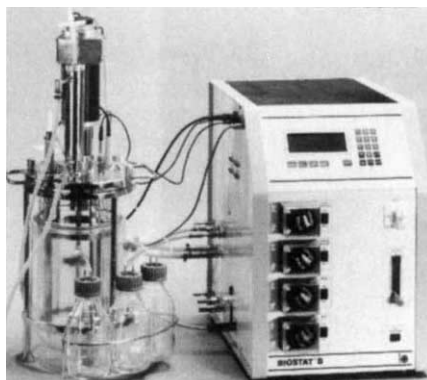
New from Unisyn Technologies is the MAB Pharm kit for producing and purifying research-scale quantities of monoclonal antibody (*Reader Service No. 109*). Three of the company's products make up the kit: the Micro Mouse hollow-fibre bioreactor cell culture system, Hybrid Grow basal medium specialized for growing hybridomas in Cell-Pharm systems, and AvidChrom for monoclonal antibodies, a purification kit specially formulated for purifying monoclonals. The Cell-Pharm Micro Mouse comes presterilized and preassembled, and can fit into an incubator. All that is needed to begin operation is a peristaltic pump. The Micro Mouse can be used to produce from 50 to 1,000 mg of monoclonal antibody per month. Hybrid Grow specialized basal medium for producing hybridomas contains elevated glucose and glutamine levels in order to deliver nutrients more efficiently to tissue-dense, hollow-fibre hybridoma cultures. The



MAB production/purification system.

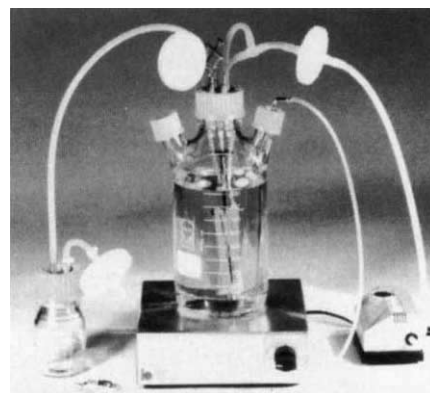
product is free of any bovine products and, as such, is intended to simplify downstream processing. Completing the package is the AvidChrom all-in-one purification kit for purifying monoclonals. Like the Micro Mouse, AvidChrom for monoclonal antibodies is designed for easy scale-up from single-use cartridges to reusable, prepacked radial flow columns.

Biostat B is B. Braun Biotech's compact laboratory **fermentation system** especially designed to meet the requirements of small-scale fermentation applications (*Reader Service No. 110*). Available with culture



Biostat B fermentation system.

vessels of 2, 5 or 10 l working volume, Biostat B can be operated batchwise and continuously, and is equipped with a measurement and control system for pH, pO_2 , antifoam, level and substrate. Process values can be registered on a chart recorder via six recorder outputs. All process values can be sent to a printer via a built-in RS232 serial

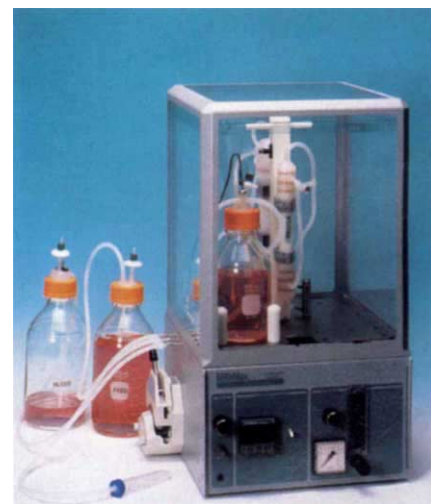


SuperSpinner culture system.

interface. For connection to a host computer, the Biostat B is equipped with an RS422 serial interface. Also available from the company, the SuperSpinner cultivation system for cell cultures to be operated in a CO_2 incubator cabinet. The system is equipped with a membrane aeration agitator and is designed for the production of small amounts of biomolecules (for example, monoclonal antibodies), as well as for the preparation of cell inoculum. The reaction vessel consists of a Duran glass flask with a working volume of 500 ml up to 1,000 ml.

Dr F. Messi Cell Culture Technologies offers **protein-free cell culture systems** in which newly selected Chinese Hamster Ovary (CHO) cells can be grown in both static and agitated culture without any supplementation of animal sera, proteins, peptides, steroids or complex additives such as hydrolysates and extracts (*Reader Service No. 111*). In mechanically agitated cultures, the company says that cell growth kinetics are comparable to those observed with serum-dependent culture processes. The cells are said to proliferate at a doubling time of 15 h, with a viability exceeding 90 per cent. The serum- and protein-independent CHO easy C cells grown in the chemically defined medium, Chomaster, may provide a useful culture system for the expression and production of recombinant proteins. According to the company, production levels higher than 100 mg l^{-1} can be achieved, depending on the protein to be expressed.

BioWhittaker's MABMAX bioreactor systems from Setec can be used in the culture of cells using hollow-fibre techniques (*Reader Service No. 112*). The Tricentric bioreactors at the heart of the MABMAX control systems have been designed to over-



MABMAX bioreactor system.

come the major limitations of hollow-fibre cell culture, which can include low cell viability, large diffusional gradients for oxygen and nutrients, frequent contamination, cell debris and lytic enzymes degrading to the product, high levels of media consumption, and poor scale-up ability. BioWhittaker says that the set-up used with the Tricentric bioreactors ensures that no cells are at a greater distance than 200 μm from oxygen and nutrients. Using this approach, stated cell densities of over 2×10^8 cells per millilitre can be achieved. The Tricentric bioreactors can be autoclaved.

■ Peripherals

The new Model DS320 Petri plate pourer and stacker is capable of processing up to 15 plates per minute, 900 plates an hour

(Reader Service No. 113). Designed as a companion unit to Jouan's line of SH series automatic agar sterilizers, the DS320 can accommodate up to 320 plates in eight stacks of 40 plates (90–95 mm × 15 mm), and can fill them unattended. The volume range can be varied from 5 ml to 99 ml, with a stated reproducibility and accuracy of ± 0.5 ml and ± 5 per cent, respectively. A covered filling zone and ultraviolet lamp are designed to protect the plates from possible external contamination. Safety features include automatic shut-off, in addition to audible and visible alarms for inverted or incomplete dishes, broken dishes, blocked elevator or pump, and interrupted power supply.

WASP is the acronym for Don Whitley Scientific's new Whitley advanced **spiral plater** (Reader Service No. 114). At the heart of the system is a microprocessor controlling all aspects of the deposition of liquid sample onto the surface of a rotating agar



Spiral plater system.

plate. Differences in height of the agar are accommodated by the special geometry of the stylus. The stylus arm dispenses liquid sample either uniformly across the plate or as a continuously decreasing volume (equivalent to a thousandfold dilution). Preprogrammed options allow liquid sample to be dispensed in three ways, utilizing up to 200 µl of sample. One special feature of the WASP system is its ability to load sample, inoculate a plate and then clean the stylus with only one keystroke. WASP has an automatic cleaning cycle, with provision for sterile water and a bleach reservoir system.

A complete line of **orbital shakers** is now available from Forma Scientific (Reader Service No. 115). All five tabletop and console models include microprocessor control with run programmes that can be defined by the user, and an independent platform monitoring system. The console modules and tabletop incubator shaker feature independent overtemperature monitoring for product protection. The console



Tabletop and console orbital shakers.

modules offer a HEPA-filtered airflow system and a stainless steel interior with coved corners. Forma has a light stability model designed specially for photosynthesis studies.

Techne has introduced a new range of MCS **biological stirrers** intended for suspension cell culture and for microcarrier beads (Reader Service No. 116). The softstart feature is said to provide controlled acceleration of the stirrer rod to the selected speed. Four stirrer sizes are available, accommodating culture vessels from 125 ml to 5,000 ml. Each unit incorporates fully variable interval stirring, as well as continuous stirring controllable up to 80 r.p.m. The user can select 'on' times between 6 s and 5 min and 'off' times between 2 min and 2 h. The motor and controls are housed within a polished stainless steel platform. All models are suitable for use in incubators/warm rooms. Two of the models will also take a specially designed water bath which maintains a constant temperature to keep the culture medium within the set limits.

A microprocessor-controlled version of LTE Scientific's Series 250 laboratory **autoclave** is now available (Reader Service No. 117). Cycles are initiated at the touch of a button. From then on, LTE says that the system operates automatically, with each stage during the cycle displayed on an illuminated 'mimic' diagram. A twin-channel printer records cycle temperatures and times, and provides a means of tracing the load by logging identification numbers and cycle parameters. Up to eight cycles can be programmed. Once these values are set and stored, a tamper-proof lockout code only allows adjustment by an authorized user.

■ Workstations

For laboratory applications requiring oxygen-free environments, Heraeus has introduced the WA6200 **anaerobic workstation** (Reader Service No. 118). Suggested uses for the workstation include culturing, inoculating and incubating oxygen-sensitive microorganisms in disciplines such as micro-

ADVERTISEMENTS

Xenopus Oocyte Injection

Singer Mk I Micromanipulator

- Oocyte Injection
- Embryo Injection
- Complete Systems

SINGER
INSTRUMENTS

"A responsibility to Science"

Tel: (+44) 984 40226 Fax: (+44) 984 41166

Singer Instrument Co. Ltd., Roadwater, Watchet, Somerset, England, TA23 0QL, UK

Reader Service No.33

Rhein Biotech

- Affordable Products -

Hansenula derived

r-HIRUDIN

Bulk Quantities • Highest Quality

HIGH YIELD = FAIR PRICE

Rhein Biotech GmbH

Phone: +49 211 709010 • Fax: +49 211 7090130

Eichsfelder Str. 11 • D-40595 Düsseldorf
Germany

Reader Service No.426

NEW AGAROSE IN THE MARKET

- METHOXYL DERIVATIZED
- LOW GELLING/MELTING

Our product is the first of its kind in the market. This Agarose derivative performs advantageously in the following ways:

- Matrix is less brittle for safer handling
 - Higher gel strength
 - Smaller exclusion limit for improved protein separation
 - Excellent resolution in PCR fragments
- Technical information and samples are available on request from:

**HISPANAGAR S.A., P.O. BOX 392,
09080 BURGOS, SPAIN**

Telephone: 34 + 47 29 85 19 Fax: 34 + 47 29 85 18

Reader Service No.5

O₂/CO₂ RESPIROMETER

Measures O₂ Consumption and CO₂ Production (BOD Rate) in Headspace of I-80 Chambers. Sensitivity 0.2 µl/h. Fully Computerized. Optional Methane Sensor. Chambers headspace volume: 50 ml to 10 l.

Applications:

Bioremediation of soil or water, bacterial or fungal growth, oxidation of food, respiration of insects, animals or plants.

For More Information Contact:

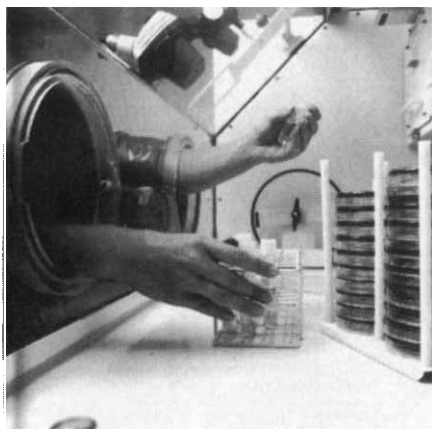
COLUMBUS INSTRUMENTS

P.O. Box 44049

Columbus, OH 43204-0049 USA

Ph: (614) 276-0861 Fax: (614) 276-0529

Reader Service No. 4



Oxygen-free work environments.

biology, genetics, pharmacology and medical microbiology. Two sealed openings at the front enable users to work within the controlled environment without gloves. As a safety feature, the openings can be 'locked' to prevent unauthorized access. A plexiglass front panel is designed to provide an unobstructed view. The WA6200 incorporates a microprocessor-controlled anaerobic incubator for working at temperatures up to 70 °C, an anaerobic airlock with a travelling carriage for the safe loading of samples. It can accommodate up to 200 Petri dishes, and can be supplied with a range of accessories, including catalyst cartridges, stacking racks, a lighting unit, a microscope adapter and a connector hook-up for gases.

The new range of **vented workstations** from Flow Sciences has been designed specifically for use in the trimming of tissue and for small animal necropsy (*Reader Service No. 119*). The company's two-foot model fits on a standard laboratory bench and incorporates air foils, turning vanes and a rear plenum designed to prevent turbulence, while at the same time maintaining a safe, steady air flow. A four-foot model is also available. The worksurfaces are 'dished' and will contain a 1.5-l spill: 3 l in the larger model. The



Vented workstation for animal necropsy.

units are available with a fan filter housing with a variable-speed fan — HEPA or catalytic filters.

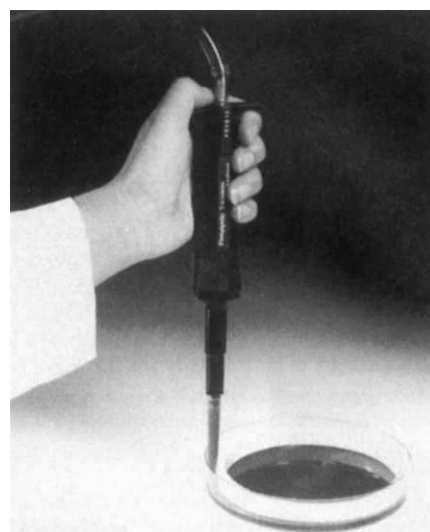
Reagents

Sigma has introduced four new **leukaemia inhibitory factor products**: human recombinant LIF and mouse recombinant LIF cytokines, as well as goat anti-human LIF and goat anti-mouse LIF polyclonal antibodies (*Reader Service No. 120*). Leukaemia inhibitory factor is a multifunctional glycoprotein that induces macrophage differentiation. It plays an important role along with IL-6 and G-CSF in the regulation of early haematopoietic stem cells. Sigma's growth factors and antibodies to growth factors are supplied with a data sheet, which includes references and bioassay data. Also available — new serum-free medium kits designed to aid growth and maintenance of targeted cell types. The fibroblast medium kit is suitable for maintaining fibroblast cultures or establishing primary cultures. Proliferation of normal human epidermal keratinocytes is supported by the keratinocyte medium kit. A kit is also available for growing normal human melanocytes. All three specialty media are serum-free. The endothelial medium kit supports the growth of endothelial cells from a variety of species, — human, bovine, rabbit, dog, mouse and rat — in a low-serum (2 per cent) environment.

For **virus inactivation in fetal calf serum** PAA Biologics offers Viralex fetal calf serum (*Reader Service No. 121*). In addition to BVD, IBR and PI-3, the company says that Viralex has been demonstrated to be effective at inactivating Foot and Mouth, Reo, Parvo, SV40 viruses and mycoplasma. The process is said to offer several benefits over conventional methods of virus inactivation (including gamma irradiation): there is little or no effect on proteins and cell growth rates for most types of cells, and no reduction in the shelf life of the serum.

Liquid transfer

The Finn timer Vacuum Pipette from Labsystems is a **vacuum pipettor** especially designed for tissue culturing techniques or where as-



Vacupette vacuum pipettor.

pirating by vacuum suction is required (*Reader Service No. 122*). The Vacupette works by attaching a tube from the top of the pipette to a vacuum system. The liquid is aspirated by a disposable tip through the pipettor to a waste container. The device can be autoclaved.

Two new **injectors** for the injection and/or transfer of 'micro' volumes of liquids into living cells can now be obtained from Eppendorf (*Reader Service No. 123*). The Transjector Model 5246, designed to replace the company's Model 5242 injector, now features an on-board pressure supply, a new remote keypad controller and dual-pressure outlets so as to allow two separate substances to be injected using this one device. Transjector Plus has all the features of the basic model but, in addition, has the ability to transfer 'micro' volumes of material from one cell to another. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

ADVERTISEMENTS

OLIGONUCLEOTIDES

Synthesised, purified and shipped in 3 days

£2.50 per base

(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit

King's College School of Medicine & Dentistry
Tel: 071 346 3126 Fax: 071-733 3877

Reader Service No.10

DNA SEQUENCING

Using an ABI 373 Automated Sequencing System

£60/sample

- PCR products - single stranded DNA
- M13 inserts - double stranded DNA
- plasmid inserts - solid phase sequencing
- 300-400 bases of sequence

Please contact us for more information and a preparation protocol

Molecular Medicine Unit

King's College School of Medicine & Dentistry
Tel: 071 346 3126 Fax: 071-733 3877

Reader Service No.11